JOURNAL OF ATMOSPHERIC AND TERRESTRIAL PHYSICS

EDITORIAL ADVISORY BOARD

H. O. G. ALFVÉN–STOCKHOLM · **J. BARTELS**–GÖTTINGEN · **W. J. G. BEYNON**–ABERYSTWYTH · **P. M. S. BLACKETT**–LONDON · **H. G. BOOKER**–ITHACA, N.Y. **E. G. BOWEN**–SYDNEY · **SIR DAVID BRUNT**–LONDON · **R. BUREAU**–BAGNEUX **S. CHAPMAN**–ALASKA · **J. EGEDAL**–COPENHAGEN · **V. F. HESS**–NEW YORK · **E. O. HULBURT**–WASHINGTON, D.C. · **H. ISRAEL**–AACHEN · **D. K. MENZEL**–CAMBRIDGE, MASS. · **S. K. MITRA**–CALCUTTA · **M. NICOLET**–UCCLE · **E. H. NORINDER**–UPSALA · **N. V. PUSHKOV**–MOSCOW · **J. A. RATCLIFFE**–CAMBRIDGE · **O. E. H. RYDBECK**–GOTHENBURG · **J. C. W. SCOTT**–OTTAWA · **SIR GEORGE SIMPSON**–LONDON · **H. SOLBERG**–OSLO · **J. VELDKAMP**–DE BILT · **M. WALDMEIER**–ZÜRICH · **J. ZENNECK**–ALTHEGNENBERG

EDITOR-IN-CHIEF

SIR EDWARD APPLETON, F.R.S.

EDINBURGH

VOLUME 16

1959

PERGAMON PRESS

NEW YORK • LONDON • PARIS • LOS ANGELES

OL.
16
959

J. Atmosph. Terr. Phys.

VOL.
16
1959

CONTENTS

Nos. 1/2

G. F. Walton: Calculation of the vertical distribution of atmospheric ozone . . 1

G. S. Kent: High frequency fading observed on the 40 Mc/s wave radiated from artificial satellite 1957 α. 10

P. C. Clemmow and M. A. Johnson: A contribution to the theory of the motion of weak irregularities in the ionosphere 21

N. F. Barber and D. D. Crombie: V.L.F. reflections from the ionosphere in the presence of a transverse magnetic field 37

L. G. H. Huxley: A discussion of the motion in nitrogen of free electrons with small energies, with reference to the ionosphere 46

J. M. Malville: Type *B* aurorae in the Antarctic 59

Walter Becker: New methods and some results concerning true ionospheric height calculation 67

P. Bandyopadhyay: The early morning *E2*-layer and some evidence of pre-sunrise *F*-layer splitting 84

W. L. Price: Anomalous ionospheric reflection during solar eclipses 93

J. D. Whitehead: Sudden changes in the virtual height of radio waves reflected from the *E*-region of the ionosphere 99

M. D. Vickers: The effect of the *F1*-layer on the calculation of the height of the *F2*-layer 103

C. A. Onwumechilli and N. S. Alexander: Variations in the geomagnetic field at Ibadan, Nigeria—I. Solar variations 106

C. A. Onwumechilli and N. S. Alexander: Variations in the geomagnetic field at Ibadan, Nigeria—II. Lunar and luni-solar variations in the *H* and *Z* 115

E. Woyk: The fraction of radio waves by a spherical ionized layer 124

A. J. Ferraro and J. J. Gibbons: Polarization computations by means of the Multislab approximation 136

B. N. Singh and R. L. Ram: Rhythmic fading of short-wave radio signals 145

A. G. McNamara and D. W. R. McKinley: The effect of trail irregularities on the interpretation of meteor echoes 156

F. Mariani: A new theoretical model of the composite *F*-layer 160

J. E. Kupperian, Jr., E. T. Byram and H. Friedman: Molecular oxygen densities in the mesosphere at Fort Churchill 174

Research Notes:

W. Dieminger: Transient fine structure of the *E*-layer 179

D. Stranz: Solar activity and the altitude of the tropopause near the equator . 180

S. Datta, P. Bandyopadhyay and R. N. Datta: Ionospheric observations on the *F*-region during the solar eclipse on April 19, 1958 182

B. J. Fulton, L. E. Petrie and W. S. P. Ward: Transient modes of high-frequency radio wave propagation across the auroral zone 185

G. A. M. King: The ionospheric *E*-layer at Cape Hallett 186

K. Davies: The effect of the earth's magnetic field on m.u.f. calculations . . 187

N. C. Gerson: Annual distribution of sporadic *E* 189

S. Evans: Horizontal movements of visual auroral features 191

V. Marasigan: Single and double inflections on the *F*-trace 193

W. R. Piggott: The gyrofrequency in the *E* layer above Slough, England . . 197

K-Indices of geomagnetic activity (December 1958–June 1959) 199

Nos. 3/4

K. C. Clark and A. E. Belon: Spectroscopic observations of the great aurora of 10 February 1958—I. Abnormal vibration of N_2^+ 205

A. E. Belon and K. C. Clark: Spectroscopic observations of the great aurora of 10 February 1958—II. Unusual atomic features 220

M. L. Meeks and J. C. James: Meteor radiant distributions and the radio-echo rates observed by forward scatter 228

E. A. Lytle and D. M. Hunten: The ratio of sodium to potassium in the upper atmosphere 236

J. F. Noxon, A. W. Harrison and A. Vallance Jones: The infra-red spectrum of the night airglow from 1·4 μ to 4·0 μ 246

G. Kvifte: Auroral and nightglow observations at Ås, Norway 252
H. K. Paetzold: The fluctuations of the acceleration of satellites 259
R. L. Ingraham: Note on the tidal theory of the S_q magnetic field 263
C. A. Onwumechilli: The relation between H- and Z-variations near the equatorial electrojet 274
J. D. Whitehead: The absorption of short radio waves in the D-, E- and F-regions of the ionosphere 283
J. A. Fejer and R. W. Vice: An investigation of the ionospheric D-region 291
J. A. Fejer and R. W. Vice: The use of full-wave solutions in the interpretation of ionospheric absorption measurements 307
J. Egedal: On a pseudo-iterating method for reduction of observations of earth-tide for effects from tides of the seas, with an application to observations made at Bergen (Norway) 1934 318
K. Bibl, A. Paul und K. Rawer: Die Frequenzabhängigkeit der ionosphärischen Absorption 324
D. M. Schlapp: Some measurements of collision frequency in the E-region of the ionosphere 340
G. V. Groves: A theory for determining upper-atmosphere winds from radio observations on meteor trails 344
J. M. Kelso: The determination of the electron density in interplanetary space . . 357
J. A. Gledhill: The effects of a solar eclipse on a stratified ionosphere 360
J. A. Gledhill: The behaviour of the ionosphere over Cape Town and Johannesburg during the annular solar eclipse of 25 December 1954 367
R. N. Singh and S. R. Khastgir: Study of "winds" in the F-region of the ionosphere during the unusual days in the IGY-calendar 376
J. S. Greenhow and E. L. Neufeld: Turbulence at altitudes of 80–100 km and its effects on long-duration meteor echoes 384
Research Notes:
D. R. Bates and M. R. C. McDowell: Escape of helium 393
W. H. Ward: E-region winds 394
Book Reviews 396

VOL. 16 1959

AUTHOR INDEX

(B.R.) denotes a book review

Alexander N. S. 106, 115
Appleton E. V. (B.R.'s) 396, 397, 398, 400
Bandyopadhyay P. 84, 182
Barber N. F. 37
Bates D. R. 393
Becker W. 67
Belon A. E. 205, 220
Bibl K. 324
Booker H. G. (B.R.) 398
Byram E. T. 174
Clark K. C. 205, 220
Clemmow P. C. 21
Crombie D. D. 37
Datta R. N. 182
Datta S. 182
Davies K. 187
Dieminger W. 179
Egedal J. 318
Evans S. 191
Fejer J. A. 291, 307
Ferraro A. J. 136
Friedman H. 174
Fulton B. J. 185
Gerson N. C. 189
Gibbons J. J. 136
Gledhill J. A. 360, 367
Greenhow J. S. 384
Groves G. V. 344
Harrison A. W. 246
Hunten D. M. 236
Huxley L. G. H. 46
Ingraham R. L. 263
James J. C. 228
Johnson M. A. 21
Jones A. Vallance 246
Kelso J. M. 357
Kent G. S. 10
Khastgir S. R. 376
King G. A. M. 186
Kupperian J. E. Jr. 174
Kvifte G. 252
Lapwood E. R. (B.R.) 399
Lytle E. A. 236
McDowell M. R. C. 393
McKinley D. W. R. 156
McNamara A. G. 156
Malville J. M. 59
Marasigan V. 197
Mariani F. 160
Martin D. (B.R.) 396
Meeks M. L. 228
Neufeld E. L. 384
Noxon J. F. 246
Onwumechilli C. A. 106, 115, 274
Paetzold H. K. 259
Paul A. 324
Petrie L. E. 185
Piggot W. R. 197
Price W. L. 93
Ram R. L. 145
Rawer K. 324
Runcorn S. K. (B.R.) 397
Schlapp D. M. 340
Singh B. N. 145
Singh R. N. 376
Stranz D. 180
Vice R. W. 291, 307
Vickers M. D. 103
Walton G. F. 1
Ward W. H. 394
Ward W. S. P. 185
Whitehead J. D. 99, 283
Woyk E. 124

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 1 to 9. Pergamon Press Ltd. Printed in Northern Ireland

Calculation of the vertical distribution of atmospheric ozone

G. F. WALTON
Physics Department, Queen's College, Dundee

(*Received* 1 *September* 1958)

Abstract—The expressions which arise in the umkehr method A of determining the vertical distribution of ozone are evaluated and depicted as a set of curves. By means of these curves zenith sky measurements with the Dobson spectrophotometer can be quickly converted into ozone distribution for any value of the total ozone amount. The method is applied to observational data from Ahmedabad and Oxford.

1. INTRODUCTION

STATIONS of the ozone network arranged by the International Ozone Commission are equipped with a Dobson spectrophotometer, with which observations can be made of the light received from the clear zenith sky. Such observations for different solar zenith angles are usually known as "umkehr observations," from which information about the vertical distribution of ozone can be obtained. Two methods, A and B, of calculating the vertical distribution at Arosa were devised by GÖTZ *et al.* (1934), and these methods have been used elsewhere by TØNSBERG and LANGLO (1944), KARANDIKAR and RAMANATHAN (1949) and LARSEN (1955). Generally there has not been uniformity in the adoption of data such as absorption and scattering coefficients, and in the division of the atmosphere into sections, with the result that the vertical distributions obtained in the past are not strictly comparable. In order to facilitate the comparison of future results, standard methods of calculation have been prepared; method A by WALTON (1957), and method B by RAMANATHAN and DAVE (1957). Both methods entail lengthy computations.

The purpose of the present work is to describe a method of finding the vertical distribution of ozone which does not require lengthy computations. By using a set of curves derived from the tables of the standard method A, the calculation of the vertical distribution is greatly simplified. The method should be useful in routine reductions of umkehr measurements.

2. THEORY AND CONSTRUCTION OF THE CURVES

Umkehr measurements with the Dobson spectrophotometer are expressed as the number N defined as follows:

$$N = \text{const.} - 100 \log I/I' \tag{1}$$

where I and I' denote the intensities at wavelengths λ and λ' of scattered light received from the zenith sky. The constant in equation (1) can be eliminated if we write

$$N(Z_2) - N(Z_1) = 100[\log I/I' \text{ at } Z_1 - \log I/I' \text{ at } Z_2] \tag{2}$$

where Z_1 and Z_2 signify two different zenith angles of the sun. On the theoretical

side, considering only primary scattering, we have the following expression (WALTON, 1957):

$$\frac{I}{I'} = \frac{\beta}{\beta'} \frac{10^{-\alpha x}(A_0 + \sum_1^4 A_r 10^{-\alpha} Y_r)}{10^{-\alpha' x}(A_0' + \sum_1^4 A_r' 10^{-\alpha'} Y_r)}, \tag{3}$$

where α, β and α', β' denote the decimal absorption and scattering coefficients at wavelengths λ and λ' respectively. Y_r is expressed in terms of the sectional amounts of ozone assigned according to the scheme shown in Table 1.

Table 1. Division of the atmosphere into sections

Section no. (r)	Height limits (km)	Assigned amount of ozone (cm)
0	54–∞	0
1	48–54	$0{\cdot}057x_1$
	42–48	$0{\cdot}204x_1$
	36–42	$0{\cdot}739x_1$
2	24–36	x_2
3	12–24	$x - x_1 - x_2 - u$
4	0–12	u

The symbol x denotes the total ozone amount and is known from direct sun measurements with the spectrophotometer, u is assumed to be $0{\cdot}085x$; x_1 and x_2 remain to be determined from two equations of the form (3) combined with observational data of the form (2). Thus if we put

$$Q(Z) = \frac{A_0 + \sum_1^4 A_r 10^{-\alpha} Y_r}{A_0' + \sum_1^4 A_r' 10^{-\alpha'} Y_r} \tag{4}$$

then $$\log Q(Z_1)/Q(Z_2) = \log I/I' \text{ at } Z_1 - \log I/I' \text{ at } Z_2 \tag{5}$$

and, from equations (2) and (5), we have

$$100 \log Q(Z_1)/Q(Z_2) = N(Z_2) - N(Z_1) = \Delta N \tag{6}$$

which equates the theoretical and observational expressions. The ozone amounts x_1 and x_2 are found in any particular case by solving two equations of the form (6) using combinations of the zenith angles 60°, 80° and 86·5°.

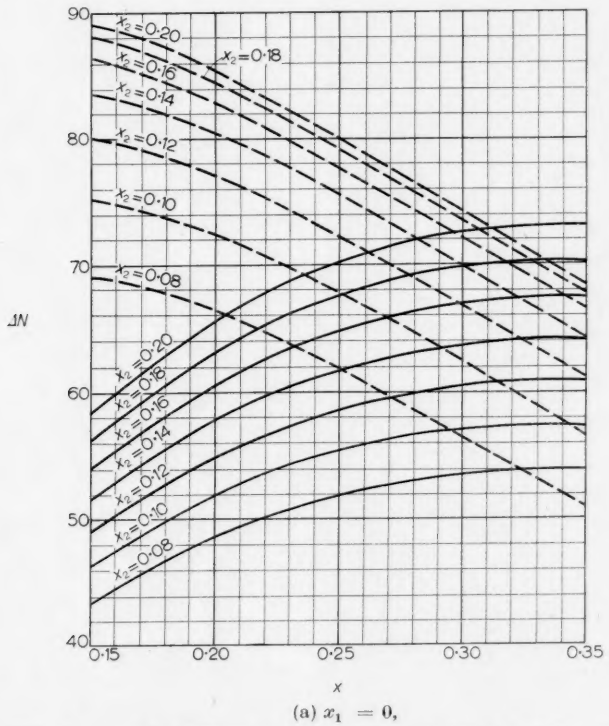
The solution of equations (6) for different values of x is facilitated by the use of the (ΔN, x) curves shown in Figs. 1(a)–(e). These have been constructed for the C wavelengths 3112/3323 Å by evaluating $\log Q(60°)/Q(80°)$ and $\log Q(60°)/Q(86{\cdot}5°)$ for five values of x, namely 0·15, 0·20, 0·25, 0·30 and 0·35 cm, five values of x_1, namely, 0, 0·01, 0·02, 0·03 and 0·04 cm and seven values of x_2 covering a suitable range in each case. A few additional points have been obtained to confirm the

VOL. 16 1959

shapes of some of the curves. In these calculations the absorption coefficients of the standard method A have been adopted; these are the "Ny" and "Choong" coefficients, not the "Vigroux" coefficients now in force (DOBSON, 1957). The constants A_r and A_r' which occur in equations (3) and (4) depend on the scattering coefficients and those of the standard method A have been adopted. The use of the curves is exemplified in Section 3.

3. APPLICATION TO DATA FROM AHMEDABAD

RAMANATHAN and DAVE (1957) have processed umkehr data obtained at Ahmedabad on 16 October 1956 ($x = 0{\cdot}192$ cm) by method B, and we here treat the same data by method A. From their Table 9 we get $\Delta N = N(80°) - N(60°) =$



(a) $x_1 = 0$,

Fig. 1. (ΔN, x) curves for different x_1 and x_2 (cm). Full lines: $\Delta N = N(80°) - N(60°)$. Dashed lines: $\Delta N = N(86{\cdot}5°) - N(60°)$. Figure continued on pp. 4 and 5.

60·2 and $\Delta N = N(86{\cdot}5°) - N(60°) = 94{\cdot}3$. Applying the provisional corrections for secondary scattering given in their Table 7 these values became 55·7 and 88·3 respectively.

Entering Figs. 1(a)–(e) with the values of x and ΔN, and interpolating from the curves, we get two sets of values of x_1 and x_2 which when plotted lead to the (x_1, x_2) curves shown in Fig. 2. The co-ordinates of the point of intersection of these

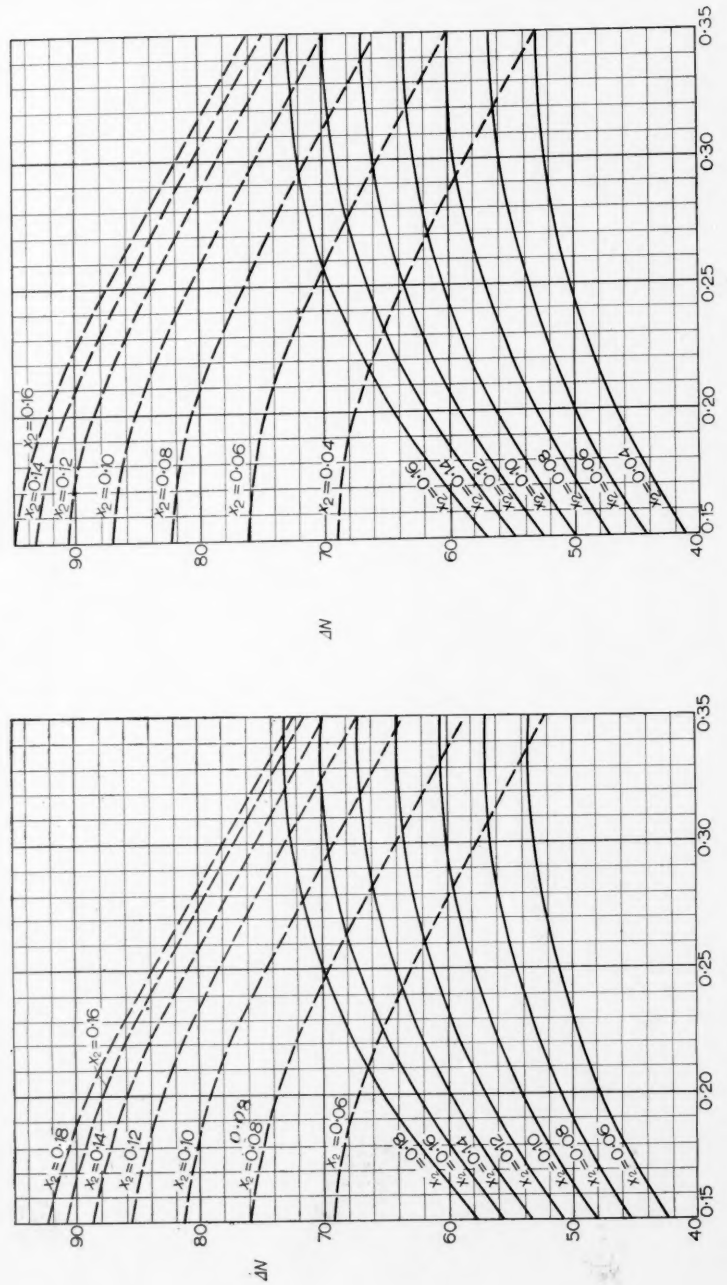
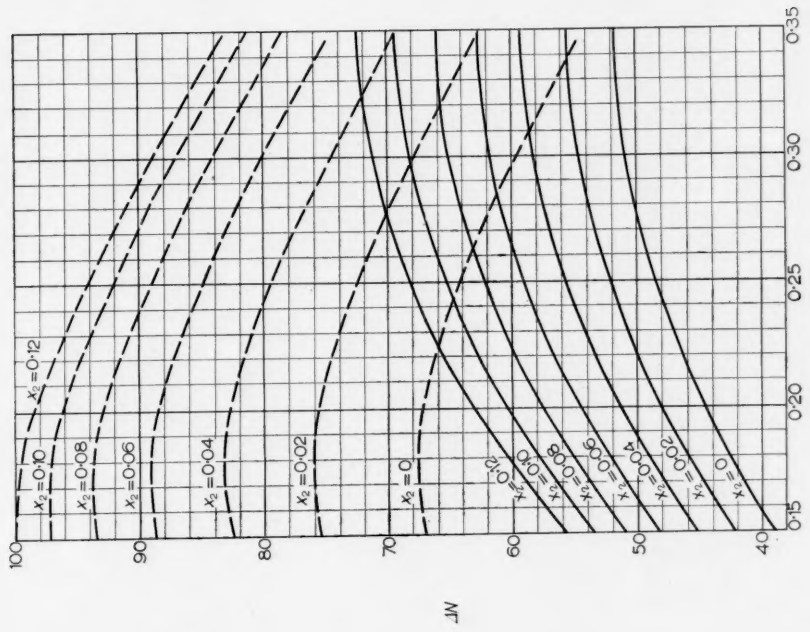
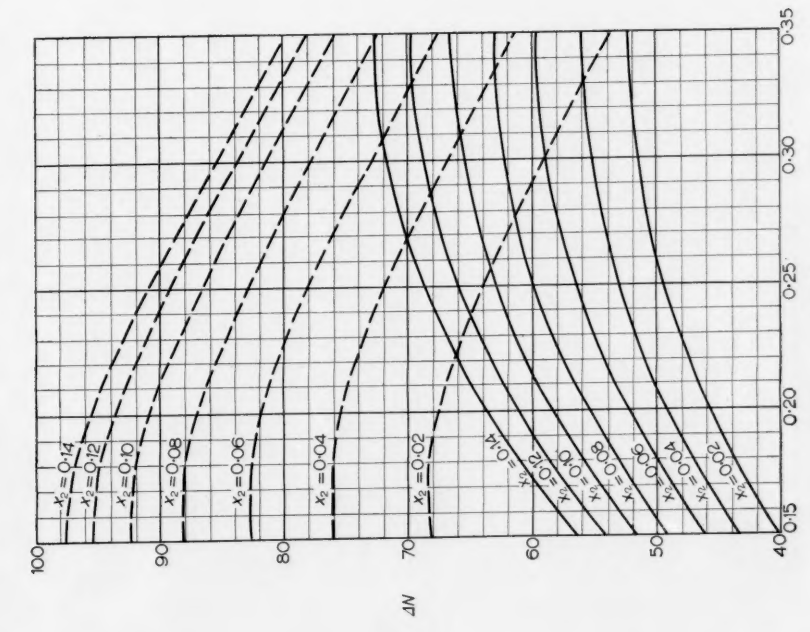


Fig. 1 (*Contd.*)

VOL.
16
1959



(e) $x_1 = 0{\cdot}04$.



(d) $x_1 = 0{\cdot}03$,

Fig. 1 (*Contd.*)

curves are the solutions of equations (6). From Fig. 2 we have $x_1 = 0{\cdot}030$ cm and $x_2 = 0{\cdot}121$ cm from the uncorrected curves, and $x_1 = 0{\cdot}027$ cm and $x_2 = 0{\cdot}093$ cm from the corrected curves. The resultant vertical distributions of ozone are shown in Figs. 3 and 4, together with the distributions given by method B.

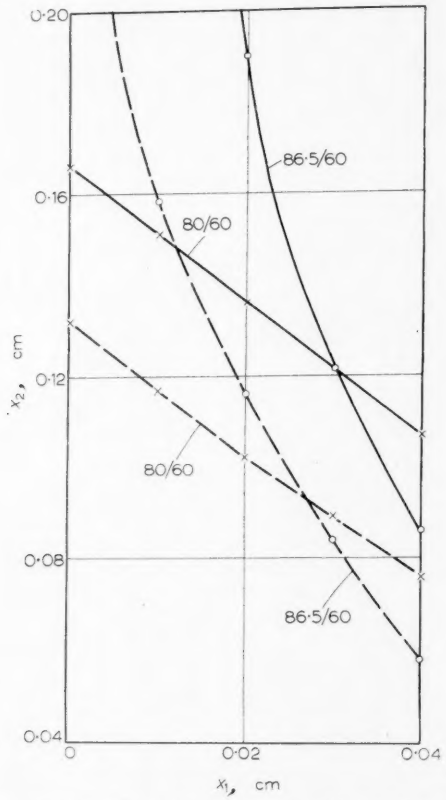


Fig. 2. (x_1, x_2) curves. The broken curves are obtained from data corrected for secondary scattering.

Method B necessarily gives a more detailed picture below 36 km than method A. Above 36 km the two methods give almost identical results, suggesting that the trial distributions required in the calculations of method B could be based on the results given by the standard method A. A comparison of Figs. 3 and 4 shows that the effect of correcting for secondary scattering is to decrease the determined amount of ozone in the atmosphere above 24 km and to increase the determined amount of ozone below this level. This effect is shown by each method.

The accuracy of the results depends on the accuracy of the umkehr measurements. The Dobson spectrophotometer is estimated to measure N to 0·5, i.e. to $\pm 0{\cdot}25$, hence the error in ΔN is $\pm 0{\cdot}25\sqrt{2}$. From the uncorrected Ahmedabad data we find that $x_1 = 0{\cdot}030 \pm 0{\cdot}002$ cm and $x_2 = 0{\cdot}121 \pm 0{\cdot}005$ cm, whence $x - x_1 - x_2 - u = 0{\cdot}025 \pm 0{\cdot}003$ cm, since $u = 0{\cdot}016$ cm. Thus the ozone

VOL. 16 1959

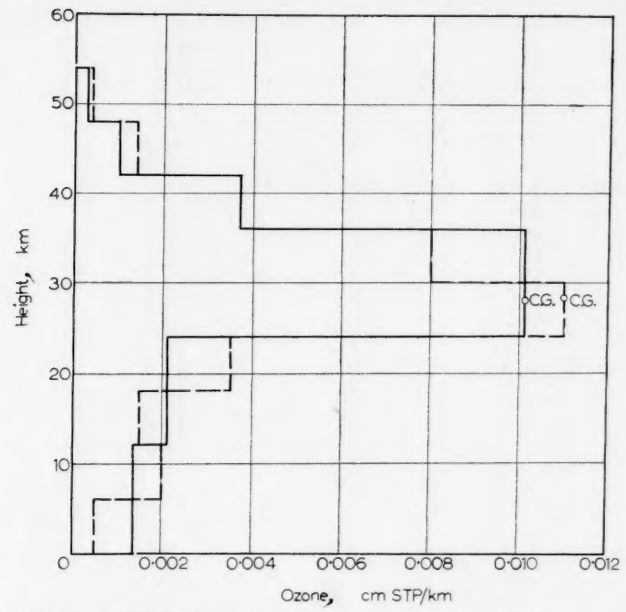


Fig. 3. Ozone distribution at Ahmedabad on 16 October 1956 (uncorrected):
——— method A; – – – – method B.

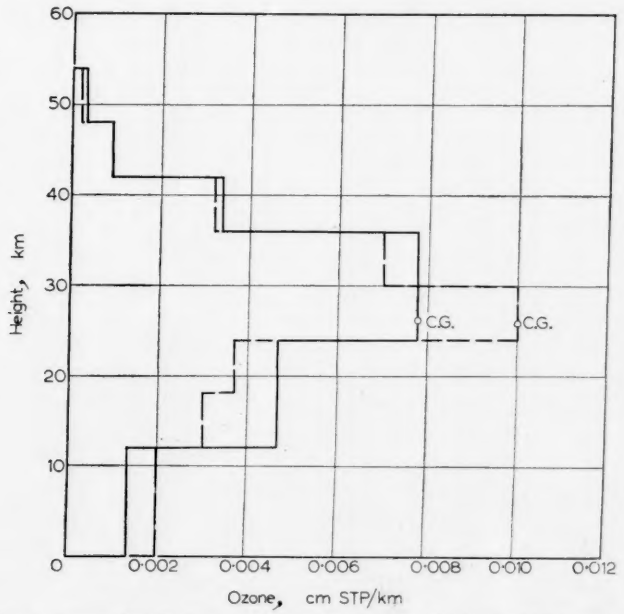


Fig. 4. Ozone distribution at Ahmedabad on 16 October 1956 (corrected):
——— method A; – – – – method B.

amounts in the 36–54 km, 24–36 km and 12–24 km sections are subject to errors of ± 7, ± 4 and ± 12 per cent respectively. The errors in these sectional amounts due to interpolation of the $(\Delta N, x)$ curves and determination of the point of intersection of the (x_1, x_2) curves are estimated to be ± 3, ± 2 and ± 8 per cent respectively.

4. Application to Data from Oxford

Several umkehr measurements obtained at Oxford by Dobson are shown in Table 2 together with the ozone distributions calculated by the process already described. The umkehr data refer to the C wavelengths, and the value of x is an average amount for the day as determined from direct sun observations.

The heights of the centre of mass of ozone shown in Table 2 have been calculated taking into account the subdivision of the 36–54 km section into the three sections given in Table 1. The results in general show that the ozone amounts in

Table 2. Uncorrected umkehr data from Oxford and calculated sectional amounts of ozone (cm) at S.T.P.

Date	21/4/53	13/5/50	9/3/53	1/6/51	24/3/51	31/5/51	30/3/51
x(cm)	0·222	0·257	0·276	0·280	0·286	0·290	0·335
$N(80°) - N(60°)$	59·9	67·5	66·8	65·0	64·7	66·1	69·6
$N(86{\cdot}5°) - N(60°)$	86·5	83·5	84·0	81·5	79·9	81·9	79·2
36–54 km	0·024	0·014	0·024	0·023	0·021	0·026	0·028
24–36 km	0·105	0·154	0·126	0·114	0·114	0·112	0·124
12–24 km	0·074	0·067	0·103	0·119	0·127	0·127	0·155
0–12 km	0·019	0·022	0·023	0·024	0·024	0·025	0·028
Height of c.g. (km)	25·1	25·4	24·5	23·7	23·4	23·5	23·5

all sections tend to increase as x increases, the increase being particularly marked in the 12–24 km section. The height of the centre of mass tends to decrease slightly as x increases. A similar descent of the centre of mass with increasing ozone has been found elsewhere by other workers. The height of the centre of mass at Oxford (52°N) is slightly higher than that found at Spitzbergen (78°N) by Larsen (1955) and slightly lower than that found at Delhi (28°N) and Poona (19°N) by Karandikar and Ramanathan (1949), but a strict comparison is not possible because the earlier results were not obtained by standard methods.

Table 3 shows the vertical distributions calculated from the data of Table 2, to which the same provisional corrections for secondary scattering have been applied as in Section 3. The correction for secondary scattering has the same general effect on the Oxford distributions as on the Ahmedabad distribution; the average

VOL. 16 1959

Table 3. Corrected amounts of ozone (cm) at S.T.P.

x(cm)	0·222	0·257	0·276	0·280	0·286	0·290	0·335
36–54 km	0·023	0·006	0·018	0·019	0·018	0·021	0·020
24–36 km	0·077	0·137	0·106	0·093	0·091	0·093	0·110
12–24 km	0·103	0·092	0·129	0·144	0·153	0·151	0·177
0–12 km	0·019	0·222	0·023	0·024	0·024	0·025	0·028
Height of c.g. (km)	23·5	23·7	23·1	22·5	22·3	22·5	22·3

effect is to decrease the ozone amounts in the 36–54 km and 24–36 km sections by 0·005 cm and 0·020 cm respectively, and to increase the amount in the 12–24 km section by 0·025 cm.

A few of the distributions are of particular interest. Between 24 March 1951 and 30 March 1951 a rise in ozone is shown in each section, the increase being greatest in the 12–24 km section. This section generally contains the maximum of ozone (cm/km), an exception being the distribution of 13 May 1950, which shows a pronounced maximum in the region of 24–36 km.

5. Conclusion

Although priority is being given to the C wavelengths, umkehr measurements are sometimes made using the A, B and D wavelengths. For these wavelengths (ΔN, x) curves could be constructed, but investigations show that, for the A and B wavelengths, $\log Q(60°)/Q(86{\cdot}5°)$ is a double-valued function of x_2. Hence interpolation of the (ΔN, x) curves would be difficult. This difficulty could probably be removed in the case of the A wavelengths by using $\log Q(80°)/Q(86{\cdot}5°)$ instead of $\log Q(60°)/Q(86{\cdot}5°)$, but for the B wavelengths it could only be removed by using zenith angles other than 80° and 86·5°. This difficulty does not arise in the case of the D wavelengths, but the angle of intersection between the (x_1, x_2) curves is found to be very small; consequently the solved values of x_1 and x_2 would be less accurately determined than in the case of the C wavelengths.

Acknowledgements—The author is indebted to Dr. G. M. B. Dobson for providing observational data and to Mrs. N. M. Stott for help in the calculation of some of the (ΔN, x) curves. Thanks are also due to Sir Charles Normand for his interest in this work and for several stimulating discussions.

References

Dobson G. M. B.	1957	*Ann. IGY* **5,** 46.
Götz F. W. P., Meetham A. R. and Dobson G. M. B.	1934	*Proc. Roy. Soc.* A**145,** 416.
Karandikar R. V. and Ramanathan K. R.	1949	*Proc. Indian Acad. Sci.* A**29,** 330.
Larsen S. H. H.	1955	*J. Atmosph. Terr. Phys.* **6,** 46.
Ramanathan K. R. and Dave J. V.	1957	*Ann. IGY* **5,** 23.
Tønsberg E. and Langlo K.	1944	*Geofys. Publ.* **13,** 12.
Walton G. F.	1957	*Ann. IGY* **5,** 9.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 10 to 20. Pergamon Press Ltd. Printed in Northern Ireland

High frequency fading observed on the 40 Mc/s wave radiated from artificial satellite 1957α

G. S. KENT
Cavendish Laboratory, Cambridge

(*Received* 30 *July* 1958)

Abstract—A study has been made of high frequency fading observed on the 40 Mc/s wave radiated from the first Russian satellite (1957α). This fading had a frequency of a few cycles per second and it is concluded that it was produced by the passage of the waves through irregularities in the *F*-region of the ionosphere. The size and spatial distribution of these irregularities has been examined, and it is found that they are about 1 km in size and that they occur at heights above about 250 km. They usually occur at latitudes greater than 50°N and are more concentrated towards the north. It is likely that their density is controlled by magnetic rather than geographic latitude and there is a tendency for them to be grouped into large areas. These have dimensions of several hundred kilometres and are elongated along the parallels of latitude. Comparison is made of the properties of these irregularities with those believed to be responsible for radio-stellar scintillation and spread-*F* echoes.

1. INTRODUCTION

THE intensity of the wave received from an earth satellite is found to fluctuate and other workers have shown that there are periodic components in these fluctuations which can be ascribed to (a) the rotation of the satellite and (b) the Faraday rotation of the plane of polarization as the wave passes through the anisotropic ionosphere (Staff of the Mullard Radio Astronomy Observatory, Cambridge, 1957; AITCHISON and WEEKES, 1958). On records of some passages of the satellite there is in addition a more rapid irregular fluctuation which will be referred to here as the high frequency component of the fading. In this paper it will be suggested that this component results from the occurrence of large-scale irregularities of size about 1 km in certain parts of the ionosphere.

VOL. 16 1959

2. THE OBSERVATIONS

Observations were made on the 40 Mc/s wave radiated from the first Russian satellite. The signal was recorded on nearly all the transits visible from Cambridge between 12 and 24 October 1957. These all occurred between about 1900 and 0700 GMT and the high frequency component of the fading was observed on the majority of these transits.

The receiver used had a band-width of about 120 c/s and a cascade first stage to give a low noise figure. The aerial consisted of a half-wave dipole mounted at the top of a 100 ft high mast and strong signals were received whenever the satellite was above the horizon. The receiver included automatic frequency control; a frequency error meter and a manual tuning control allowed the signal to be followed accurately as the doppler frequency shift occurred. Recording was made on a pen-recorder with a time constant of about $\frac{1}{2}$ sec; this was a little too slow to follow the faster fluctuations in amplitude and a correction had to be applied to find the true depth of the fading.

Fig. 1 shows a typical record. The rapid irregular amplitude fluctuations with a period of less than 1 sec represent the high frequency component of the fading. These ceased at about 0320 GMT and can be clearly distinguished from the much

slower fading on the right of the figure. The latter is attributed to rotation of the satellite, to Faraday rotation of the plane of polarization of the wave and to the passage of the satellite through the maxima and minima of the reception pattern of the receiving aerial.

To describe the high frequency component of the fading a fluctuation index S is defined as:

$$S = \frac{\text{standard deviation of the amplitude}}{\text{mean amplitude of the signal}}$$

This quantity varies from 0 for no fading to 0·52 for fading with a Rayleigh

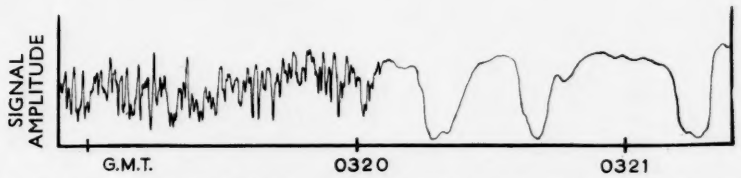


Fig. 1. Record of the 40 Mc/s signal received from the first Russian satellite showing the transition from high frequency to much slower fading.

probability distribution. The pen recorder did not respond fully to the higher fading frequencies and to correct for this the measured value of S was multiplied by a correction factor derived from a knowledge of the frequency response of the recorder. The magnitude of the factor did not vary very much and was generally less than two.

3. The Effect of Ionospheric Irregularities

If the electron density N in the ionosphere is irregularly distributed so that, over a region having linear dimensions of order l, it departs from the mean by ΔN, the amplitude and phase of a wave arriving at the ground will vary in a way that has been studied by other workers (Hewish, 1952; Bramley, 1951; Ratcliffe, 1956). If the source, or the ionospheric irregularities, move relative to the observer on the ground the amplitude and phase of the received wave will fluctuate. When the source of radiation is a radio star the relative motion of the source and observer is of less importance than the motion of the ionospheric irregularities and the fading is caused by the ionospheric movements. If the same ionospheric movements are supposed to occur in the irregularities responsible for the high frequency component of the fading observed on the signal received from a satellite then it can be shown that they are negligible compared with the motion of the satellite. The fading observed therefore corresponds to the motion of the satellite past the ionospheric irregularities and if its velocity is known the size l of the irregularities can be determined. It will be shown in Sections 5 and 6 that this estimate of l is consistent with that deduced from observations on the fluctuations of the signals from radio stars.

Previous workers (Hewish, 1952) have shown that if in its passage through the ionosphere the wave encounters on an average n irregularities and if it travels a

distance z in free space after leaving the ionosphere the fluctuation index S varies as:

$$S \propto \Delta n \, . \, n^{1/2} \, . \, z \tag{1}$$

If the irregularities occur within a limited range of heights and if the satellite is above them then n is proportional to sec i where i is the angle between the wave normal and the vertical. If the irregularities are distributed uniformly above a certain limiting height h_0 then n is proportional not only to sec i but also to the distance of the satellite above h_0.

4. The Occurrence of the Ionospheric Irregularities

4.1. *Introduction*

From what follows it will appear that the occurrence of ionospheric irregularities of sufficient importance to produce noticeable fading may depend upon the following factors:

(a) Geographical position.

(b) Height.

(c) Local time.

(d) Degree of disturbance in the ionosphere.

In this section we shall discuss the relation between the occurrence of the high speed component of the fading and the position of the satellite in its orbit.

VOL. 16 1959

4.2. *The observed occurrence of rapid fading*

Fig. 2 shows a map of Europe with the satellite paths for the night of 13–14 October (Staff of the Mullard Radio Astronomy Observatory, Cambridge, 1957; Staff of the Royal Aircraft Establishment, Farnborough, 1957). The continuous sections of the curves shown are those parts of the orbit from which signals were received. The local time at Cambridge is marked for one point on each orbit together with the direction in which the satellite was travelling.

The height of the satellite varied over the latitudes visible from Cambridge as shown in Fig. 3, so that in the evening it was observed beneath the F-region at heights of 200–300 km and in the morning it was observed above the F-region at heights of 350–500 km. The local time at the satellite for 13–14 October is also shown in Fig. 3. Over the period of observation the plane of the satellite's orbit was precessing slowly. This movement was however small and on any single night it could be considered to be fixed in space, so that a given position in the orbit corresponded to a given local time.

In Figs. 2 and 3 lines are drawn across the satellites tracks to indicate the positions of the satellite when rapid fading was observed. In Fig. 2 the lines correspond to intervals of 15 sec and their lengths are proportional to the fluctuation index S when the satellite was at these positions. In Fig. 3 their lengths are proportional to the mean fluctuation index averaged over all the transits on this night.

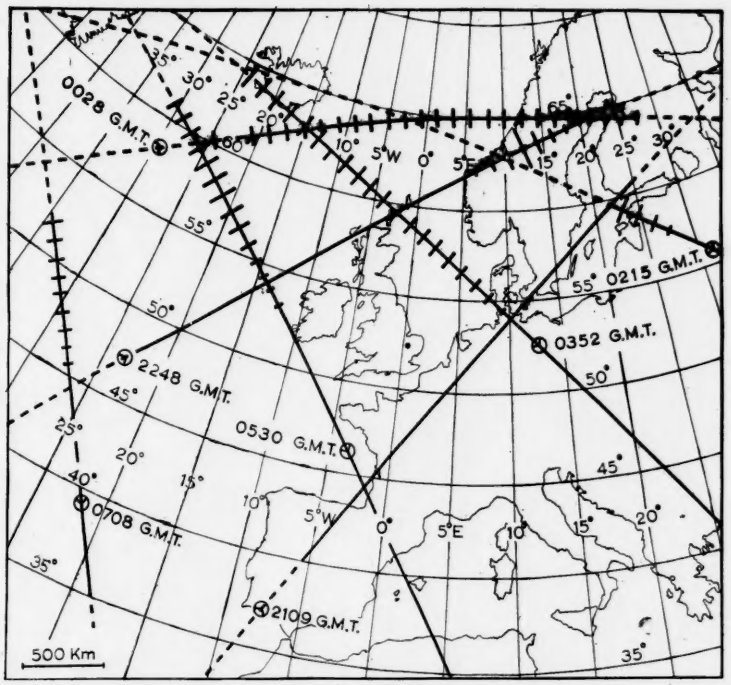


Fig. 2. Map showing the satellite tracks on 13–14 October. The parts of the orbit from which signals were received are shown by solid lines, the remaining parts by dashed lines. The short lines perpendicular to the tracks show the position of the satellite when the high frequency component of the fading was observed. They are spaced at 15 sec intervals and have their length proportional to the fluctuation index.

The main facts which emerge from the figures are the following:

(1) The high frequency component of the fading was observed only when the satellite was at the most northern part of the orbit. In addition when it lay to the west of Cambridge the fading occurred at lower latitudes than when it lay to the east. This is discussed in Section 4.4.

(2) Less fading was observed when the satellite was travelling north than when it was travelling south. This is discussed in Section 4.3.

(3) The fluctuation index did not change monotonically with position. In the transits observed at 2248 and 0530 GMT for example the satellite appeared to be passing through regions of irregularities with varying density. This is discussed in Section 4.5.

4.3. *The height distribution of the irregularities*

Fig. 3 shows the variation of the height, the local time and the fluctuation index with the latitude of the satellite on 13–14 October. The results summarized in this figure could be explained by supposing either that the irregularities producing the rapid fading occur only in a limited region of space, that they occur only during a limited interval of time, or both. Fortunately the slow precession of the orbit caused the relation between local time and position in the orbit to vary by about

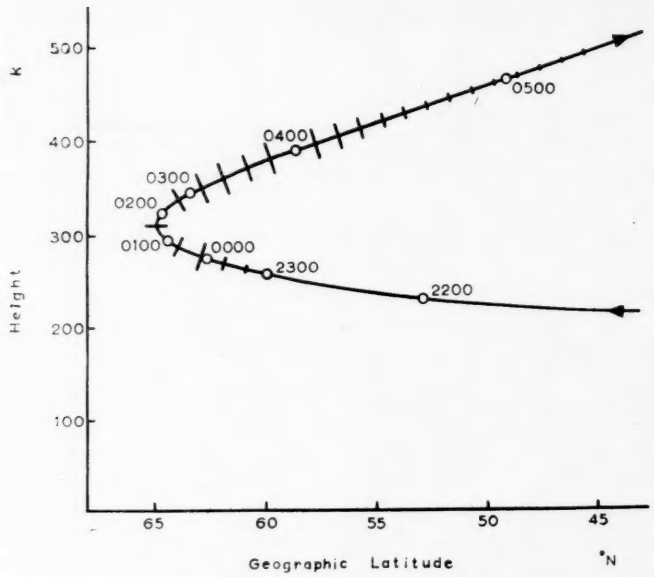


Fig. 3. Diagram showing the variation of height and local time at the satellite with latitude on 13–14 October. The short lines across the track have their length proportional to the mean fluctuation index.

VOL.
16
1959

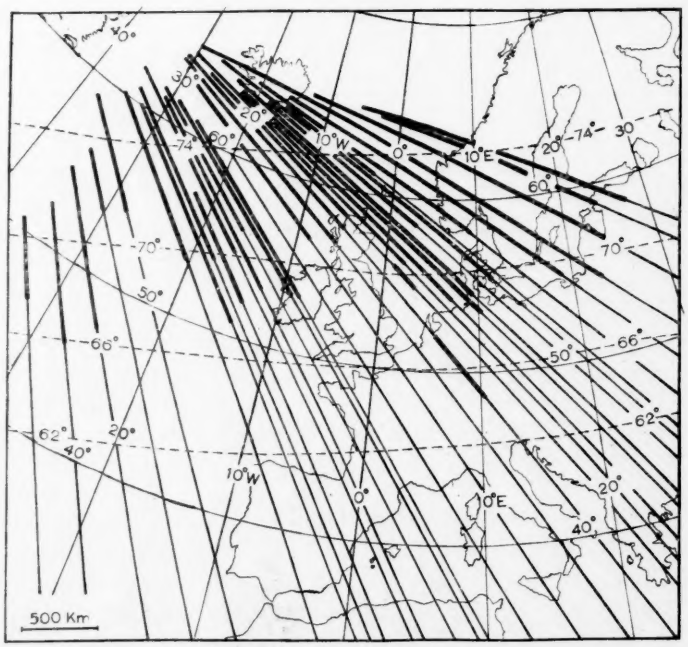


Fig. 4. Map showing all the transits that were observed when the satellite was travelling south. The position of the satellite when the high frequency component of the fading was observed is shown by a broadening of the line representing the track.

3 hr over the 12 days of the observations. By making use of this variation it was possible to show that both the appearance and disappearance of the fluctuations were determined primarily by the height and latitude of the satellite and not by the local time.

From the position of the satellite when the fluctuations first appeared as shown in Fig. 3 it was concluded that the irregularities did not occur below 250 km and rarely below 270 km. No upper limit can be given to their height but it is probable that a major proportion lies below 325 km since the fluctuation index did not usually increase after the satellite had risen to this height.

4.4. *Geographical distribution*

In the previous sections it was noticed that the high frequency component of the fading disappeared on the south-going journey of the satellite at about latitude 50°N. This occurred on every night as shown in Fig. 4 in which all the south-going transits are represented with a thickened line to indicate where the rapid fluctuations were observed.

It was stated earlier that it was possible to show that the disappearance of the

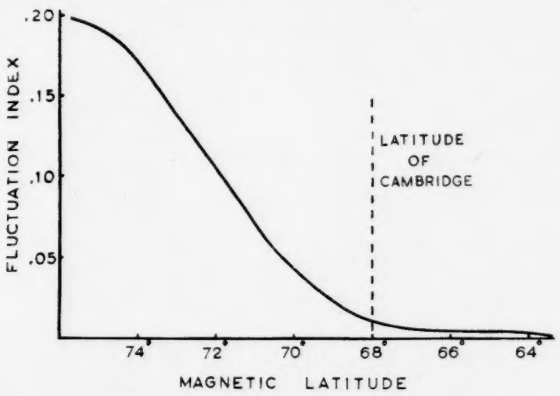


Fig. 5. Variation of the fluctuation index with magnetic latitude.

fluctuations was not determined by local time. It is therefore concluded that the irregularities causing the fading did not occur south of about 50°N. In Fig. 4 it can be seen that when the satellite lay to the west of Cambridge the fluctuations in amplitude occurred further south than when it lay to the east. A statistical analysis shows that this difference is significant and it appears probable that the density of the irregularities causing the fading is controlled more by magnetic than by geographic latitude.

The way in which the mean fluctuation index varies with magnetic latitude is shown in Fig. 5. In Section 3 we saw that the fluctuation index depends upon the angle that the wave normal makes with the vertical and the distance that the wave travels after leaving the ionosphere. It is not possible, from the present results, to decide how far the observed fluctuation index depends upon these factors and how far it is caused by an increase in the irregular nature of the ionosphere towards the north.

4.5. *The concentration of the irregularities into limited regions*

It is shown in Fig. 2 that the fluctuation index on the transits at 2248 and 0530 GMT were most marked in two separate parts of the track. To show this phenomenon more clearly the fluctuation index for the transits observed on 22–23 October is plotted in Fig. 6 as a function of the distance along the track. It is seen that the irregularities are concentrated into regions which extend over several thousands of kilometres.

In different transits the satellite is observed when its track makes different angles with the parallels of latitude. It is thus possible in principle to measure the

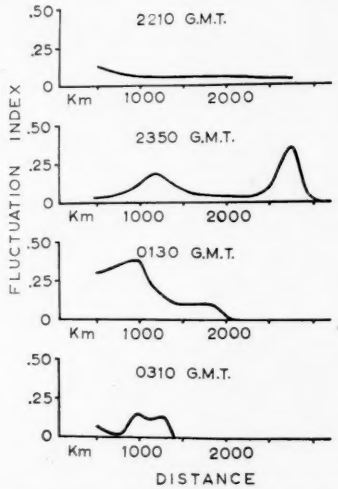


Fig. 6. Diagram showing the fluctuation index as a function of distance along the satellite tracks for the transits observed on 22–23 October.

size of these regions and to determine their shape. Analysis of the observations is complicated and involves the fact that the intensity of the irregularities increases towards the north and also the possibility that the size of the regions into which the irregularities are concentrated may change with latitude.

A detailed analysis shows with some certainty that these regions are elongated along the magnetic parallels of latitude and that their dimensions are of the order of 1000 km along and between 200 and 400 km across these parallels. It is possible but less certain that the size of these regions increases towards the north.

5. The Size of the Irregularities

Although most of the records were made with an instrument which was too slow-moving to reproduce the high speed component of the fading accurately, there was one night on which a recorder with a response time of about 0·01 sec was used. It is considered that the fading was faithfully portrayed on this occasion and part of the record has been analysed to give the auto-correlogram shown in Fig. 7(a). This auto-correlogram may be used in the following way to make deductions about the probable size of the irregularities.

VOL. 16 1959

Suppose in Fig. 7(b) that the satellite is at a point A at a height h, that the irregularities causing the amplitude fluctuations occur in a thin layer at a height h' and that the signal is observed at a point B on the ground. Let $\rho_A(\tau)$ be the time auto-correlation function of the amplitude $A(t)$ as observed at B and $\rho_f(\xi)$ the space auto-correlation function of the complex amplitude of the wave after emerging from the layer. If (a) the wave after emerging from the layer is randomly phased; (b) the scattering angles are small; (c) the probability distribution of $A(t)$ is of the "displaced Gauss" form, then it can be shown that (RATCLIFFE, 1956):

$$\rho_A(\tau) \simeq \rho_f(\xi)$$

if $\xi = v\tau\left(\frac{h'}{h}\right)$ and v is the velocity of the satellite.

When the section of record corresponding to the auto-correlation function in Fig. 7(a) was taken the height of the satellite was 450 km. Let us suppose that the

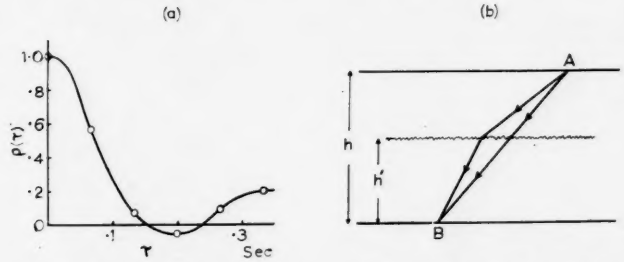


Fig. 7. (a) Auto-correlogram of the fading observed at 0300 GMT on 25 October. (b) Diagram illustrating the scattering geometry. The signal is radiated from the satellite at A and received at a point B on the ground after passing through a layer of irregularities at a height h'.

irregularities occur in a thin layer at a height of 300 km.

Then $h'/h = 0{\cdot}67$

$$v = 7 \text{ km/sec}$$

and $$\rho_A(\tau) = 0{\cdot}5 \text{ at } \tau = 0{\cdot}08 \text{ sec}$$

therefore $$\rho_f(\xi) = 0{\cdot}5 \text{ at } \xi = 0{\cdot}4 \text{ km}.$$

The scale of the complex amplitude variations after emergence from the layer is therefore about $\frac{1}{2}$ km. If in addition it is assumed that the phase changes introduced by the irregularities are less than 1 rad then the size of the irregularities will be of the same order as the size of the fluctuations in the complex amplitude pattern i.e. about $\frac{1}{2}$ km to the point where the correlation has fallen to one-half.

6. Relation to Other Results

It has already been mentioned that ionospheric irregularities in the F-region have been suggested as the cause of the scintillation of radio stars. They have also been suggested as the cause of the spread-F echoes observed when ionospheric sounders are used. In this section we discuss the relation between our results and the results of studying these other phenomena.

6.1. *The size of the irregularities*

In Section 5 the size of the irregularities was found to be about $\frac{1}{2}$ km to the point where the correlation had fallen to one-half. There have been several estimates of the size of the irregularities causing radio star scintillations (HEWISH, 1952; LITTLE and MAXWELL, 1951) and SPENCER (1955) has shown that they are elongated along the earth's magnetic field and have dimensions of a few kilometres between maxima. The value obtained above agrees well with this.

6.2. *The height of the irregularities*

There have been few estimates of the heights at which the irregularities causing radio-stellar scintillations occur. HEWISH (1952) by a consideration of the diffraction pattern formed on the ground fixed their height at about 400 km. BRIGGS

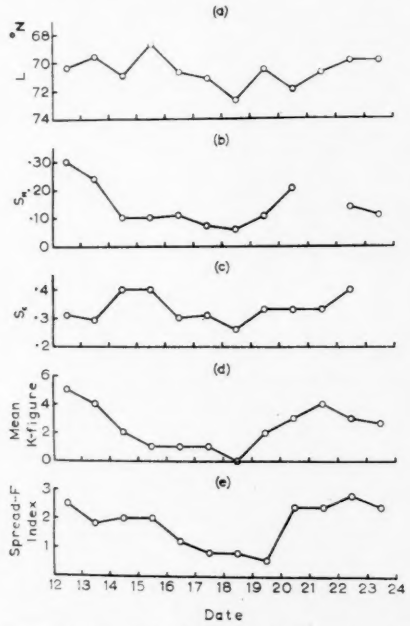


Fig. 8. Variation of characteristics of the fading and of the ionosphere over the period of observation, 12–24 October 1957. (a) Southern extent L of the irregularities causing the fading, in magnetic latitude. (b) Mean fluctuation index S_M of the fading observed on the signal from the satellite. (c) Fluctuation index S_C observed on the waves from the radio-stellar source in Cassiopeia. (d) Magnetic K-figure observed at Hartland, Devon. (e) Spread-F index observed at Inverness.

(1958a) by comparison of scintillation and spread-F placed it at 300 km. The results given in Section 4.3 suggest that a large proportion of the irregularities lies between 270 and 320 km in good agreement with the latter value.

6.3. *Variability*

Fig. 8 shows the way in which the simultaneously recorded characteristics of the fluctuations and of the ionosphere varied over the period of observation. Fig. 8(a)

VOL. 16 1959

shows a function L derived from the records which represents the southern extent of the irregularities causing the fluctuations, plotted as magnetic latitude. Fig. 8(b) shows the mean fluctuation index on each night. Fig. 8(c) shows the fluctuation index observed at Cambridge (magnetic latitude 68°N) for the radio-stellar source in Cassiopeia. Fig. 8(d) shows the magnetic K-figure recorded at Hartland, Devon (magnetic latitude 67°N). Fig. 8(e) shows the spread-F index as observed at Inverness (magnetic latitude 72°N). This is proportional to the amount of spreading on the $h'(f)$ records and varies from 0 for no spreading to 3 for so much spreading that the traces for the ordinary and extraordinary rays are completely joined.

There can be seen to be considerable daily variation in all the quantities and while those shown in Fig. 8(b), (d) and (e) behave similarly it is not possible to say whether this is significant.

6.4. *Geographical location*

From observations on a single radio star it is usually not possible to observe whether the density of irregularities causing scintillation varies with latitude. Observations have been made (Mullard Radio Astronomy Observatory, Cambridge, private communication) on two radio stars on several days during January 1958, the waves from the two sources passing through the F-region of the ionosphere at points separated by about 600 km in a north–south direction. It was found that while no scintillation was seen on the source to the south of Cambridge, strong scintillations were seen a few hours later when the signal from the source to the north of Cambridge was observed. It has also been shown that (Briggs, 1958b) more spread-F echoes are observed at Inverness (magnetic latitude 72°N) than at Slough (magnetic latitude 67°N). Both these results indicate that the density of the irregularities causing radio-stellar scintillations and spread-F echoes increases with latitude.

It is known that the irregularities causing radio-stellar scintillations are controlled by the earth's magnetic field. Spencer (1955) has shown that the individual irregularities are elongated along the direction of the field. Briggs (1958a) by a comparison of ionospheric records taken at Inverness, Oslo and Slough has concluded that the irregularities responsible for spread-F echoes occur in large concentrations and these concentrations are probably related to the magnetic latitude.

In Section 4.4 we saw that the density of the irregularities causing fluctuations in the signal from the satellite increased with latitude from Cambridge and was probably dependent upon magnetic rather than geographic latitude. This may be compared with the other less direct evidence as outlined above which suggests the same result.

6.5. *Concentration into regions*

The only previous work on the size of the regions over which radio-stellar scintillation and spread-F echoes occur is that of Briggs (1958a). From a comparison of the two he concluded that the irregularities causing them were concentrated into areas which were elongated along the parallels of latitude. The

dimensions of the areas deduced were about 450 km across the parallels of latitude and 1000 km along them. These values are in good agreement with those in Section 4.5 where it was concluded that the irregularities occurred in regions with dimensions of between 200 and 400 km across the parallels of latitude and about 1000 km along them.

Acknowledgements—The work described above was carried out as part of the programme of the Radio Research Board. The paper is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research. The author would like to thank Dr. B. H. BRIGGS for much helpful advice and discussion, and colleagues in the Laboratory for assistance in making the observations.

REFERENCES

AITCHISON G. J. and WEEKES K.		To be published.
BRAMLEY E. N.	1951	*Proc. Instn. Elect. Engrs.* **1,** 98, 19.
BRIGGS B. H.	1958a	*J. Atmosph. Terr. Phys.* **12,** 34.
BRIGGS B. H.	1958b	*J. Atmosph. Terr. Phys.* **12,** 89.
HEWISH A.	1952	*Proc. Roy. Soc.* A **214,** 494.
LITTLE C. G. and MAXWELL A.	1951	*Phil. Mag.* **42,** 267.
RATCLIFFE J. A.	1956	*Rep. Progr. Phys.* **19,** 188.
SPENCER M.	1955	*Proc. Phys. Soc. Lond.* B **68,** 493.
Staff of the Mullard Radio Astronomy Observatory, Cambridge.	1957	*Nature, Lond.* **180,** 879.
Staff of the Royal Aircraft Establishment, Farnborough.	1957	*Nature, Lond.* **180,** 937.

VOL. 16 1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 21 to 36. Pergamon Press Ltd. Printed in Northern Ireland

A contribution to the theory of the motion of weak irregularities in the ionosphere

P. C. Clemmow and M. A. Johnson
A.S.R.E., Portsdown, Cosham, Hants.

(*Received* 12 *September* 1958)

Abstract—The motion of weak irregularities in an otherwise homogeneous slightly ionized gas under the influence of electrostatic and magnetostatic fields is investigated. The governing equations are set out and the dispersion equation for harmonic plane wave solutions is obtained and solved. It is shown that, to a good approximation for ionospheric applications, and with the neglect of diffusion, a plane wave travels with negligible attenuation with a velocity which is (a) proportional to the magnitude of the electrostatic field, (b) independent of the wavelength, (c) dependent on the direction of the wave-normal. As a particular consequence, any weak two-dimensional irregularity which is parallel to the direction of the magnetostatic field travels unchanged; so also does any one-dimensional irregularity, and in this case the analysis is unaffected by its strength. The effects of diffusion are examined.

1. Introduction

A major aspect of current ionospheric research is the investigation of the movement of irregularities of ionization density. The actual conditions in the ionosphere are complicated and only partially known, but the phenomena suggest certain idealized theoretical problems the solutions to which are likely to be relevant to any interpretation of the experimental results. The authors have examined, by various means, one such problem. Its formulation is given below, preceded by a few comments which draw attention to its more distinctive features.

For the purposes of description it is necessary to refer both to the ionization and to that part of the air which is not ionized. In what follows the latter is called the "un-ionized air"; admittedly the expression is inelegant, but it avoids the ambiguity of "neutral air", which might be thought to include neutral ionization.

The main point to emphasize, perhaps, is the necessity of considering the effect of electric fields. Since electric fields are believed to exist in the ionosphere, their contribution to the motivation of the charged particles must be taken into account; it may be quite as important as that associated with a wind in the un-ionized air, and indeed would appear to be all important in the F-region. Also, of course, the effect of the earth's magnetic field must be included. Another special complication arises because the ions are not bound, and their drift motion constitutes a current; the discussion is restricted in the present paper to the simplest case, in which only positive ions and electrons are present, but the ion current must be treated on the same footing as the electron current.

The idealized problem can be stated, in general terms, as follows: in the absence of irregularities there exists a region occupied by a homogeneous gas which is slightly ionized but everywhere neutral; the region is pervaded by uniform electrostatic and magnetostatic fields; if, at any instant, some irregularity appears in the ionization, what is its subsequent behaviour?

In the problem actually discussed some further assumptions are made in order to simplify the analysis. One is that the region occupied by the gas is infinite and unbounded; admittedly, the magnetic field arising from the steady current in the gas is then indeterminate, but this is immaterial, since the magnetic fields of currents are in any case neglected. Another assumption is that the total number of charged particles of each type is taken to be constant; in other words, processes of production or disappearance of ionization are ignored. Again, it is supposed that the un-ionized air remains homogeneous, even when irregularities appear in the ionization; that is, collisional interaction of the charged particles with the neutral particles is taken to have a negligible effect on the latter, so that the collision frequencies are regarded as uniform throughout the region. Collisions between ions and electrons are not introduced explicitly, but can easily be taken into account.

It should also be remarked that a wind in the un-ionized air can be allowed for merely by taking axes moving with the air and making an appropriate modification to the electrostatic field. This case, therefore, does not require separate consideration.

The nature of the problem can be indicated quite briefly. It is supposed that an initial distribution of positive ions and electrons is given, representing an irregularity in the ionization. Since the drift velocity of the positive ions differs from that of the electrons, clearly the inhomogeneity of the ionization, even if neutral to start with, almost at once results in the creation of polarization charge. The electric field associated with the polarization charge in turn exercises its control on the drift velocities of the charged particles, and the subsequent behaviour of the irregularity, which is also subject to diffusion, can only be determined by calculation.

The problem has, in fact, been treated by the authors in two distinct ways. In the first, the irregularity is assumed to be a localized region with a uniform ionization density different from that of its surroundings; the region is permitted to carry on its boundary a surface polarization charge density of zero total charge, and a solution is sought in which the entire ionization pattern persists without change of form or density. In the second, a perturbation theory is used; the deviation of the ionization density from its mean is taken to be so small that an approximation in which the governing equations become linear, and hence far more tractable, is justified.

The different character of the two approaches is evident. In the first, solutions which conform to a prescribed behaviour are sought, and there is no obvious *a priori* reason for supposing that they do or do not exist. In fact, it appears on examination that they do exist, but only if the irregularities are either one-dimensional, or two-dimensional, of elliptic cross-section and parallel to the uniform magnetostatic field. In the perturbation method, on the other hand, no restrictions are placed on the solution other than that the disturbance which it represents is weak. Some results for a problem of the former type have already been briefly reported (Clemmow *et al.* 1954), and a more detailed and extended discussion is to be given elsewhere. The present paper is concerned with the perturbation treatment.

VOL.
16
1959

Since the equations of the perturbation theory are linear it is easy to derive solutions in which the space variation is harmonic; the time variation is then also harmonic, with frequency determined by a dispersion equation. These plane wave solutions can be superposed to yield the solution appropriate to any given initial distributions of positive ions and electrons, the superposition process being specified by a spatial Fourier analysis of the initial distributions, and leading immediately to integral expressions for the distributions at any subsequent time. It happens, however, that some interesting information can be inferred directly from the harmonic plane wave solutions, without explicitly considering their combination; the results of the present paper are obtained in this way, the mathematics thus being kept at a comparatively elementary level. The more complete theory, which is certainly required, is under investigation.

For the sake of simplicity the problem is first treated without taking account of the diffusion associated with the electron and ion pressures. The governing equations are set out, their linear form derived (Section 3), and the dispersion equation for harmonic plane wave solutions obtained (Section 4). The dispersion equation is then solved for the frequency in terms of the wavelength, and the solution simplified by the introduction of certain approximations (Section 5). It is also shown that the approximate form of the solution can be obtained rather more quickly by an alternative argument, based on equating the electron and ion densities; and it is remarked, in passing, that the latter method yields a solution for any one-dimensional irregularity, not necessarily weak. After a brief consideration of the effect of diffusion (Section 7), the solution of the dispersion equation is discussed and certain conclusions drawn (Section 8). In the course of the analysis considerable simplifications are achieved by order of magnitude arguments; the numerical estimates adopted for various quantites are therefore set out in an Appendix.

What is established in the paper concerning the approximate behaviour of an irregularity may be summed up in the following way. If diffusion is neglected: (a) any one-dimensional irregularity (in this case, not necessarily weak) travels unchanged with a certain constant velocity, which depends *inter alia* on the magnitudes and directions of the electrostatic and magnetostatic fields; (b) any weak two-dimensional irregularity which is parallel to the magnetostatic field travels unchanged with an analogous constant velocity; (c) any other type of weak two-dimensional irregularity, or any weak three-dimensional irregularity, does not preserve its form, but changes in an apparently complicated way which requires further investigation. The effect of diffusion is to produce, in addition, a weakening of the irregularity with time, accompanied by a change of shape which consists mainly of an elongation along the magnetostatic field.

2. Notation

Only positive ions and electrons are considered, so for brevity the former are simply called ions:

ε = permittivity of free space.
m_i, m_e = ion, electron mass.

e = magnitude of electron charge, a positive number.

ν_i = half the frequency of ion collisions with neutral particles.

ν_e = the frequency of electron collisions with neutral particles (the factor "half" in the definition of ν_i is included to maintain symmetry between the ion equations and the electron equations; it arises from the fact that m_i is approximately equal to the neutral particle mass, whereas m_e is very much less).

$\mathbf{B}^\circ$ = unperturbed uniform magnetostatic field.

B° = magnitude of $\mathbf{B}^\circ$.

b_1, b_2, b_3 = direction cosines of $\mathbf{B}^\circ$.

$\mathbf{E}^\circ$ = unperturbed uniform electrostatic field.

E° = magnitude of $\mathbf{E}^\circ$

e_1, e_2, e_3 = direction cosines of $\mathbf{E}^\circ$.

$\mathbf{E}$ = perturbation electric field (giving total electric field $\mathbf{E}^\circ + \mathbf{E}$).

φ = perturbation electric potential.

N = unperturbed constant uniform number density of both ions and electrons.

n_i, n_e = perturbation number density of ions, electrons (giving total number density of ions, electrons $N + n_i$, $N + n_e$).

$\mathbf{v}_i^\circ, \mathbf{v}_e^\circ$ = unperturbed constant uniform drift velocity of ions, electrons.

$\mathbf{v}_i, \mathbf{v}_e$ = perturbation drift velocity of ions, electrons (giving total drift velocity of ions, electrons $\mathbf{v}_i^\circ + \mathbf{v}_i$, $\mathbf{v}_e^\circ + \mathbf{v}_e$).

k = Boltzmann's constant.

T = absolute temperature.

The analysis is expressed in terms of the ion and electron mobilities. These are, the direct mobilities:

$$d_i = \frac{1}{m_i \nu_i}, \qquad d_e = \frac{1}{m_e \nu_e}$$

the transverse mobilities:

$$t_i = \frac{\nu_i}{m_i(\nu_i^2 + \omega_i^2)}, \qquad t_e = \frac{\nu_e}{m_e(\nu_e^2 + \omega_e^2)}$$

and the Hall mobilities:

$$h_i = \frac{\omega_i}{m_i(\nu_i^2 + \omega_i^2)}, \qquad h_e = \frac{\omega_e}{m_e(\nu_e^2 + \omega_e^2)}$$

where $\omega_i = eB^\circ/m_i$, $\omega_e = eB^\circ/m_e$ are the ion, electron gyrofrequencies.

3. The Governing Equations

If ion–electron collisions, ion and electron pressures, and magnetic fields of currents are neglected, and if the un-ionized air particles have no drift motion,

VOL. 16 1959

then the drift velocities of the ions and electrons in the unperturbed state are given by:

$$m_i\nu_i\mathbf{v}_i^\circ - e(\mathbf{E}^\circ + \mathbf{v}_i^\circ \wedge \mathbf{B}^\circ) = 0 \tag{1}$$

$$m_e\nu_e\mathbf{v}_e^\circ + e(\mathbf{E}^\circ + \mathbf{v}_e^\circ \wedge \mathbf{B}^\circ) = 0 \tag{2}$$

and in the perturbed state by:

$$m_i\nu_i(\mathbf{v}_i^\circ + \mathbf{v}_i) - e[\mathbf{E}^\circ + \mathbf{E} + (\mathbf{v}_i^\circ + \mathbf{v}_i) \wedge \mathbf{B}^\circ] = 0 \tag{3}$$

$$m_e\nu_e(\mathbf{v}_e^\circ + \mathbf{v}_e) + e[\mathbf{E}^\circ + \mathbf{E} + (\mathbf{v}_e^\circ + \mathbf{v}_e) \wedge \mathbf{B}^\circ] = 0 \tag{4}$$

Subtraction then yields:

$$m_i\nu_i\mathbf{v}_i - e(\mathbf{E} + \mathbf{v}_i \wedge \mathbf{B}^\circ) = 0 \tag{5}$$

$$m_e\nu_e\mathbf{v}_e + e(\mathbf{E} + \mathbf{v}_e \wedge \mathbf{B}^\circ) = 0 \tag{6}$$

in which only the perturbation contribution to the drift velocities and electric field appears.

From equations (1) and (2), $\mathbf{v}_i^\circ$ and $\mathbf{v}_e^\circ$ are conveniently expressed in terms of $\mathbf{E}^\circ$ through the tensor relations:

$$\mathbf{v}_i^\circ = eT_i\mathbf{E}^\circ \tag{7}$$

$$\mathbf{v}_e^\circ = eT_e\mathbf{E}^\circ \tag{8}$$

where

$$T_i = \begin{bmatrix} t_i + b_1^2(d_i - t_i) & b_3h_i + b_1b_2(d_i - t_i) & -b_2h_i + b_1b_3(d_i - t_i) \\ -b_3h_i + b_2b_1(d_i - t_i) & t_i + b_2^2(d_i - t_i) & b_1h_i + b_2b_3(d_i - t_i) \\ b_2h_i + b_3b_1(d_i - t_i) & -b_1h_i + b_3b_2(d_i - t_i) & t_i + b_3^2(d_i - t_i) \end{bmatrix} \tag{9}$$

$$T_e = \begin{bmatrix} -t_e - b_1^2(d_e - t_e) & b_3h_e - b_1b_2(d_e - t_e) & -b_2h_e - b_1b_3(d_e - t_e) \\ -b_3h_e - b_2b_1(d_e - t_e) & -t_e - b_2^2(d_e - t_e) & b_1h_e - b_2b_3(d_e - t_e) \\ b_2h_e - b_3b_1(d_e - t_e) & -b_1h_e - b_3b_2(d_e - t_e) & -t_e - b_3^2(d_e - t_e) \end{bmatrix} \tag{10}$$

Likewise equations (5) and (6) give:

$$\mathbf{v}_i = eT_i\mathbf{E} \tag{11}$$

$$\mathbf{v}_e = eT_e\mathbf{E} \tag{12}$$

In the case when the un-ionized air particles have a constant uniform drift velocity $\mathbf{u}$, equations (1) to (6) are modified by subtracting $\mathbf{u}$ from the velocities in all terms containing the collision frequencies. If, then, the equations are referred to axes moving with velocity $\mathbf{u}$, they differ from (1) to (6) only in that $\mathbf{E}^\circ$ is replaced by $\mathbf{E}^\circ + \mathbf{u} \wedge \mathbf{B}^\circ$. It is therefore a simple matter to deduce results for the case $\mathbf{u} \neq 0$ directly from those for the case $\mathbf{u} = 0$, and only the latter is considered explicitly in the subsequent analysis.

Neglect of the time derivative of the magnetic field gives:

$$\text{curl } \mathbf{E} = 0 \tag{13}$$

or

$$\mathbf{E} = -\text{grad } \varphi \tag{14}$$

From (11), (12) and (14), $\mathbf{v}_i$, $\mathbf{v}_e$ and $\mathbf{E}$ are all expressible in terms of the potential φ. There are thus three unknowns, ϕ and the number densities of ions and electrons.

The three equations determining them are that relating the space charge to the electric field:

$$\operatorname{div} \mathbf{E} = \frac{e}{\varepsilon}(n_i - n_e) \tag{15}$$

and the respective conservation equations for ions and electrons,

$$\frac{\partial n_i}{\partial t} + \operatorname{div}\{(N + n_i)(\mathbf{v}_i^\circ + \mathbf{v}_i)\} = 0 \tag{16}$$

$$\frac{\partial n_e}{\partial t} + \operatorname{div}\{(N + n_e)(\mathbf{v}_e^\circ + \mathbf{v}_e)\} = 0 \tag{17}$$

Equations (16) and (17) are evidently non-linear. But the perturbation is now assumed to be so small that they can be replaced by the approximate linear equations:

$$\frac{\partial n_i}{\partial t} + \mathbf{v}_i^\circ \,.\, \operatorname{grad} n_i + N \operatorname{div} \mathbf{v}_i = 0 \tag{18}$$

$$\frac{\partial n_e}{\partial t} + \mathbf{v}_e^\circ \,.\, \operatorname{grad} n_e + N \operatorname{div} \mathbf{v}_e = 0 \tag{19}$$

The equations on which the investigation is based are therefore (15), (18) and (19), in conjunction with the relations (11), (12) and (14). In the next section solutions representing harmonic plane wave disturbances are derived.

VOL. 16 1959

4. The Dispersion Equation for Harmonic Plane Wave Disturbances

It is first postulated that space variations exist only in one direction, say that of the x-axis. It follows from equation (14) that E_x is the only non-zero component of the perturbation electric field, and it is convenient to work in terms of E_x instead of the potential φ. Equations (15), (18) and (19) are:

$$\frac{\partial E_x}{\partial x} = \frac{e}{\varepsilon}(n_i - n_e) \tag{20}$$

$$\frac{\partial n_i}{\partial t} + \mathbf{v}_{ix}^\circ \frac{\partial n_i}{\partial x} + N \frac{\partial v_{ix}}{\partial x} = 0 \tag{21}$$

$$\frac{\partial n_e}{\partial t} + \mathbf{v}_{ex}^\circ \frac{\partial n_e}{\partial x} + N \frac{\partial v_{ex}}{\partial x} = 0 \tag{22}$$

Without loss of generality $\mathbf{B}^\circ$ is taken to lie in the x-z-plane, so that $b_2 = 0$, and equations (9), (10) are:

$$T_i = \begin{bmatrix} b_1^2 d_i + b_3^2 t_i & b_3 h_i & b_1 b_3 (d_i - t_i) \\ -b_3 h_i & t_i & b_1 h_i \\ b_1 b_3 (d_i - t_i) & -b_1 h_i & b_3^2 d_i + b_1^2 t_i \end{bmatrix} \tag{23}$$

$$T_e = \begin{bmatrix} -b_1^2 d_e - b_3^2 t_e & b_3 h_e & -b_1 b_3 (d_e - t_e) \\ -b_3 h_e & -t_e & b_1 h_e \\ -b_1 b_3 (d_e - t_e) & -b_1 h_e & -b_3^2 d_e - b_1^2 t_e \end{bmatrix} \tag{24}$$

Equations (7) and (8) therefore give:

$$v_{ix}^{\circ} = e(b_1^2 d_i + b_3^2 t_i)E_x^{\circ} + eb_3 h_i E_y^{\circ} + eb_1 b_3 (d_i - t_i)E_z^{\circ} \tag{25}$$

$$v_{ex}^{\circ} = -e(b_1^2 d_e + b_3^2 t_e)E_x^{\circ} + eb_3 h_e E_y^{\circ} - eb_1 b_3 (d_e - t_e)E_z^{\circ} \tag{26}$$

It follows, of course, from equations (11) and (12), that equations (25) and (26) also hold when $\mathbf{v}_i^{\circ}$, $\mathbf{v}_e^{\circ}$, $\mathbf{E}^{\circ}$ are replaced by $\mathbf{v}_i$, $\mathbf{v}_e$, $\mathbf{E}$, respectively. But $E_y = E_z = 0$, so:

$$v_{ix} = e(b_1^2 d_i + b_3^2 t_i)E_x \tag{27}$$

$$v_{ex} = -e(b_1^2 d_e + b_3^2 t_e)E_x \tag{28}$$

The substitution of equations (27) and (28) into equations (21) and (22), and the use of equation (20), give two equations to be solved for n_i and n_e, namely:

$$\frac{\partial n_i}{\partial t} + v_{ix}^{\circ} \frac{\partial n_i}{\partial x} + \frac{Ne^2}{\varepsilon}(b_1^2 d_i + b_3^2 t_i)(n_i - n_e) = 0 \tag{29}$$

$$\frac{\partial n_e}{\partial t} + v_{ex}^{\circ} \frac{\partial n_e}{\partial x} - \frac{Ne^2}{\varepsilon}(b_1^2 d_e + b_3^2 t_e)(n_i - n_e) = 0 \tag{30}$$

Spatially periodic solutions of equations (29) and (30) are now considered. If the variation of n_i and n_e with x enters only through a factor:

$$\exp(-ikx) \tag{31}$$

the general solution for n_i and n_e as functions of t can be obtained by taking a time dependence given solely by the factor:

$$\exp(i\omega t) \tag{32}$$

and finding ω as a function of k. In fact, equations (29) and (30) become:

$$\left[\omega - kv_{ix}^{\circ} - \frac{iNe^2}{\varepsilon}(b_1^2 d_i + b_3^2 t_i)\right]n_i + \frac{iNe^2}{\varepsilon}(b_1^2 d_i + b_3^2 t_i)n_e = 0 \tag{33}$$

$$\left[\omega - kv_{ex}^{\circ} - \frac{iNe^2}{\varepsilon}(b_1^2 d_e + b_3^2 t_e)\right]n_e + \frac{iNe^2}{\varepsilon}(b_1^2 d_e + b_3^2 t_e)n_i = 0 \tag{34}$$

and the equality of the two expressions thus given for the ratio n_i/n_e leads to the dispersion equation:

$$\frac{b_1^2 d_i + b_3^2 t_i}{\omega - kv_{ix}^{\circ}} + \frac{b_1^2 d_e + b_3^2 t_e}{\omega - kv_{ex}^{\circ}} = -\frac{i\varepsilon}{Ne^2} \tag{35}$$

5. The Solution of the Dispersion Equation

The problem envisaged is that in which there is a given initial distribution of n_i and n_e. The dispersion equation (35) is therefore regarded as an equation for ω in terms of a given real value of k. It is evidently a quadratic, and yields, in general, two complex values of ω; the existence of two values arising from the possibility of specifying initially both n_i and n_e. The real part of ω is associated with the velocity of a travelling wave; the imaginary part of ω is shown below to be

positive, and hence, by virtue of the factor (32), is associated with exponential decay with time.

The solution of equation (35) for ω in terms of k can be written formally:

$$2\omega = k(v_{ix}^{\circ} + v_{ex}^{\circ}) + \frac{i}{\tau}\{1 \pm \sqrt{(1 - 2i\delta Lk - L^2k^2)}\} \tag{36}$$

where, for the sake of brevity, the time interval:

$$\tau = \frac{\varepsilon}{Ne^2\{b_1^2(d_i + d_e) + b_3^2(t_i + t_e)\}} \tag{37}$$

is introduced, together with the length:

$$L = \tau(v_{ix}^{\circ} - v_{ex}^{\circ}) \tag{38}$$

and the dimensionless quantity:

$$\delta = \frac{b_1^2(d_i - d_e) + b_3^2(t_i - t_e)}{b_1^2(d_i + d_e) + b_3^2(t_i + t_e)} \tag{39}$$

First it is shown that the imaginary part of ω is always positive. Let:

$$1 - 2i\delta Lk - L^2k^2 = (p - iq)^2 \tag{40}$$

where p and q are real. Then $q^2 = p^2 - 1 + L^2k^2$, so that $|p| \geqslant 1$ implies $|q| \geqslant |Lk|$, which is impossible since $pq = \delta Lk$ and $|\delta| < 1$. Hence $|p| < 1$, which establishes the result, since from equation (36) the imaginary part of ω is $(1 \pm p)/(2\tau)$.

VOL. 16 1959

It is, of course, easy to derive explicit expressions for p and q. But this is hardly worthwhile, since a simple approximation is available when $|Lk| \ll 1$, a condition which is likely to hold in ionospheric applications (even for a spatial wavelength $2\pi/k$ as short as 50 m the value of $|Lk|$ would never be more than 10^{-2} and usually much less; see Appendix). The two values of ω obtained by expanding the radical in equation (36) in powers of Lk are:

$$\omega_1 = \frac{i}{\tau} + Uk + O(L^2k^2) \tag{41}$$

$$\omega_2 = Vk + i\frac{1 - \delta^2}{4\tau}L^2k^2 + O(L^3k^3) \tag{42}$$

where:

$$U = \frac{(b_1^2d_i + b_3^2t_i)v_{ix}^{\circ} + (b_1^2d_e + b_3^2t_e)v_{ex}^{\circ}}{(b_1^2d_i + b_3^2t_i) + (b_1^2d_e + b_3^2t_e)} \tag{43}$$

$$V = \frac{(b_1^2d_e + b_3^2t_e)v_{ix}^{\circ} + (b_1^2d_i + b_3^2t_i)v_{ex}^{\circ}}{(b_1^2d_i + b_3^2t_i) + (b_1^2d_e + b_3^2\,t_e)}, \tag{44}$$

The presence of the decay term i/τ in equation (41) means that $\exp(i\omega_1 t)$ becomes negligible after a time interval of the order of τ. But τ is probably never greater than 2×10^{-4} sec, and usually much less (see Appendix), so that in practice ω_1 can safely be ignored; in fact, it plays a part only during the initial

rapid redistribution of the electron and ion charge densities which results in the appropriate polarization charge.

In equation (42), on the other hand, of the two terms explicitly retained, the first is associated with a wave travelling with velocity V, and the second with a decay of this wave with time. But the decay is comparatively slow, never less than several seconds and usually of the order of minutes or hours (see Appendix), and in any case is unimportant compared with the decay due to diffusion (see Section 7).

The conclusion, therefore, is that, apart from a decay with time governed by diffusion, a harmonic wave whose wavelength is greater than a few metres travels substantially without change of form with the velocity V. Since V is independent of the wavelength, it follows that this is the velocity of travel of any weak one-dimensional irregularity not containing structure finer than a few metres. But of more importance is the fact that V is dependent on the direction of propagation of the wave relative to the directions of the electrostatic and magnetostatic fields; this dependence, and some of its implications, are discussed in Section 8.

6. An Alternative Argument

The conclusion stated at the end of Section 5 might have been obtained more quickly, though perhaps less precisely, by making early in the analysis an approximation which is an important general guide in problems of this type. This approximation is based on the fact that a polarization field **E** of the sort of magnitude existing in the ionosphere is associated with only a relatively small difference in the ion and electron densities, and consists essentially in replacing equation (15) by $n_i = n_e$.*

If this is done, equations (21) and (22), with the use of (27) and (28), become:

$$\frac{\partial n}{\partial t} + v_{ix}^{\circ}\frac{\partial n}{\partial x} + Ne(b_1^2 d_i + b_3^2 t_i)\frac{\partial E_x}{\partial x} = 0 \tag{45}$$

$$\frac{\partial n}{\partial t} + v_{ex}^{\circ}\frac{\partial n}{\partial x} - Ne(b_1^2 d_e + b_3^2 t_e)\frac{\partial E_x}{\partial x} = 0 \tag{46}$$

where n is the common value of n_i and n_e. The elimination of $\partial E_x/\partial x$ then yields simply:

$$\frac{\partial n}{\partial t} + V\frac{\partial n}{\partial x} = 0 \tag{47}$$

where V is given by equation (44).

Equation (47) states that any common ion, electron density distribution that depends only on x moves with velocity V.

In terms of the analysis of Section 5, the result of setting $n_i = n_e$ and ignoring equation (15) is to make the right-hand side zero in the dispersion equation (35), which then becomes just $\omega = Vk$.

It is worth remarking, incidentally, that in the one-dimensional case this

* Stated in a possibly more familiar form in terms of the total drift current density j, the procedure neglects the displacement current and takes div $j = 0$ (see for example Cowling (1957) pp. 2 and 3).

simplifying technique still leads to equation (47) even if the non-linear continuity equations (16) and (17) are used instead of their linearized counterparts (18) and (19). For in the one-dimensional case the forms of (16) and (17) corresponding to (45) and (46) are:

$$\frac{\partial n}{\partial t} + v_{ix}\circ\frac{\partial n}{\partial x} + e(b_1^2 d_i + b_3^2 t_i)\left\{E_x\frac{\partial n}{\partial x} + (N+n)\frac{\partial E_x}{\partial x}\right\} = 0 \tag{48}$$

$$\frac{\partial n}{\partial t} + v_{ex}\circ\frac{\partial n}{\partial x} - e(b_1^2 d_e + b_3^2 t_e)\left\{E_x\frac{\partial n}{\partial x} + (N+n)\frac{\partial E_x}{\partial x}\right\} = 0 \tag{49}$$

Evidently E_x and $\partial E_x/\partial x$ can be simultaneously eliminated and equation (47) recovered.

The broad conclusion, therefore, is that, apart from the effect of diffusion, any one-dimensional irregularity, not necessarily weak, travels without change of form with the velocity V.

7. The Effect of Diffusion

It is possible to take account of the diffusion associated with the ion and electron partial pressures by adding the forces grad $p_i/(N+n_i)$, grad $p_e/(N+n_e)$ to the left-hand sides of equations (3) and (4) respectively, where p_i and p_e are the perturbation pressures (see, for example, COWLING (1957), Chapter 6). Equations (11) and (12) are then replaced by:

$$\mathbf{v}_i = T_i\left(e\mathbf{E} - \frac{\text{grad } p_i}{N+n_i}\right) \tag{50}$$

$$\mathbf{v}_e = T_e\left(e\mathbf{E} + \frac{\text{grad } p_e}{N+n_e}\right) \tag{51}$$

Since $p_i = \boldsymbol{k}Tn_i$, $p_e = \boldsymbol{k}Tn_e$, where $\boldsymbol{k}$ is Boltzmann's constant and T the temperature, the terms arising from the inclusion of the partial pressures when expressions (50) and (51) for $\mathbf{v}_i$ and $\mathbf{v}_e$ are substituted into equations (16) and (17) are respectively:

$$-\boldsymbol{k}T \text{ div } (T_i \text{ grad } n_i) \tag{52}$$

$$\boldsymbol{k}T \text{ div } (T_e \text{ grad } n_e) \tag{53}$$

If variations exist only in the x-direction, (52) and (53) are:

$$-\boldsymbol{k}T(b_1^2 d_i + b_3^2 t_i)\frac{\partial^2 n_i}{\partial x^2} \tag{54}$$

$$-\boldsymbol{k}T(b_1^2 d_e + b_3^2 t_e)\frac{\partial^2 n_e}{\partial x^2} \tag{55}$$

The effect of these terms is seen most quickly by adopting the argument of Section 6; for they appear (with $n_i = n_e = n$) simply as additions to the left-hand sides

VOL. 16 1959

of equations (48) and (49) respectively, from which it follows at once that the new equation corresponding to (47) is:

$$\frac{\partial n}{\partial t} + V\frac{\partial n}{\partial x} = D\frac{\partial^2 n}{\partial x^2} \tag{56}$$

where

$$D = \frac{2\boldsymbol{k}T(b_1^2 d_i + b_3^2 t_i)(b_1^2 d_e + b_3^2 t_e)}{(b_1^2 d_i + b_3^2 t_i) + (b_1^2 d_e + b_3^2 t_e)} \tag{57}$$

D is the "ambipolar" diffusion coefficient referring to the common motion of the electrons and ions.

For a harmonic plane wave disturbance of the type considered in Section 4, in which the time and space variations appear only in the factor:

$$\exp\{i(\omega t - kx)\}$$

(56) gives:

$$\omega = Vk + iDk^2 \tag{58}$$

so that diffusion leads to an exponential decay with time, in which the exponent is inversely proportional to the square of the spatial wavelength and depends on the direction of the wave normal relative to that of the magnetostatic field.

The numerical values given in the Appendix show that diffusive decay predominates over the decay associated solely with electromagnetic effects, discussed in Section 4, except low in the D-region, where in any case the decay due to either cause is quite negligible. On the other hand it is for the most part not rapid enough to destroy the validity of the concept of the velocity of an irregularity or to prevent its measurement by radio-methods. A noticeable effect might, however, arise in the F-region due to the smallness there of the diffusion time constant $1/(k^2D)$ when $b_3 = 0$, that is, for disturbances with wavenormals along $\mathbf{B}^\circ$; for this implies that, whatever their initial shape, irregularities could quite quickly become stretched along the earth's magnetic field. This point has been emphasized by SPENCER (1955), who gives some experimental evidence that irregularities are thus elongated.

8. DISCUSSION

Within certain limitations it has been established that, under the influence of electric and magnetic fields, a plane harmonic disturbance travels with a certain velocity which is independent of the wavelength, but does depend, *inter alia*, on the direction of the wavenormal. The limitations of the theory are now briefly considered, and then explicit expressions for the velocity displayed and discussed.

How valid is the assumption that the irregularities are weak? Little is known about their strength; the deviation of the ionization density from the mean density is perhaps generally thought to be of the order of 1 per cent, or less, of the mean density, except for certain travelling disturbances in the F-region, where the corresponding figure might be 50 per cent. As a working hypothesis which makes the problem much more tractable the assumption can hardly be severely criticized.

Again, there come immediately to mind two effects ignored in the theory as presented; those due respectively to ion–electron collisions and to the presence

of negative ions. Ion–electron collisions are likely to be important in the F-region, and it is a straightforward matter to take them into account; for example, if diffusion is neglected they result only in a modification of the direct mobilities, d_i, d_e, both of which become multiplied by the factor $\nu_e/(\nu_e + \nu)$, where ν is the ion–electron collision frequency. In principle the analysis can also be extended to allow for the presence of negative ions, though this is more of a complication since it increases by unity the degree of the dispersion equation. However, it is generally agreed that there are comparatively few negative ions in the ionosphere, and it is reasonable to disregard their effect until this opinion is revised.

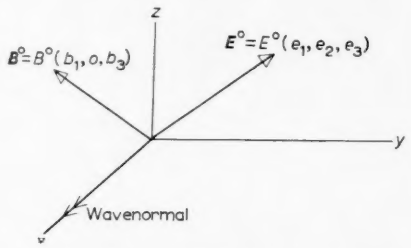


Fig. 1. The configuration to which formula (59) refers.

VOL. 16 1959

The velocity V of a plane harmonic disturbance, given by expression (44), is now examined in more detail. To see explicitly how it depends on the ionospheric parameters, the expressions (25) and (26) for $v_{ix}{}^\circ$ and $v_{ex}{}^\circ$, together with those for the mobilities in Section 2, are substituted into equation (44). After a little algebra it is found that:

$$V = \frac{e_2\left(b_1{}^2 + \dfrac{\nu_i\nu_e}{\omega_i\omega_e}\right) - e_3\left(\dfrac{\nu_i}{\omega_i} - \dfrac{\nu_e}{\omega_e}\right)}{\dfrac{\nu_i\nu_e}{\omega_i\omega_e}\left(1 + \dfrac{\nu_i\nu_e}{\omega_i\omega_e}\right) + \left(1 + \dfrac{\nu_i{}^2}{\omega_i{}^2} - \dfrac{\nu_i\nu_e}{\omega_i\omega_e} + \dfrac{\nu_e{}^2}{\omega_e{}^2}\right) b_1{}^2}\, b_3 \frac{E^\circ}{B^\circ} \tag{59}$$

where it is recalled that the direction of the wave normal is along Ox, that $\mathbf{B}^\circ$ is in the Oxz plane with direction cosines (b_1, o, b_3), and that $\mathbf{E}^\circ$ has direction cosines (e_1, e_2, e_3). The configuration is shown in Fig. 1.

A useful alternative form to (59) is obtained by taking axes $O\xi\eta\zeta$ fixed with respect to the directions of $\mathbf{B}^\circ$ and $\mathbf{E}^\circ$, allowing the direction cosines (l_1, l_2, l_3) of the wavenormal to be arbitrary. With the configuration of Fig. 2, in which the ξ-axis is along $\mathbf{B}^\circ$, and $\mathbf{E}^\circ = E^\circ(\varepsilon_1, o, \varepsilon_3)$ lies in the $O\xi\zeta$ plane, it is found that:

$$V = \frac{\varepsilon_3\left(l_1{}^2 + \dfrac{\nu_i\nu_e}{\omega_i\omega_e}\right) l_2 + \left(\dfrac{\nu_i}{\omega_i} - \dfrac{\nu_e}{\omega_e}\right) l_1[l_1(l_1\varepsilon_1 + l_3\varepsilon_3) - \varepsilon_1]}{\dfrac{\nu_i\nu_e}{\omega_i\omega_e}\left(1 + \dfrac{\nu_i\nu_e}{\omega_i\omega_e}\right) + \left(1 + \dfrac{\nu_i{}^2}{\omega_i{}^2} - \dfrac{\nu_i\nu_e}{\omega_i\omega_e} + \dfrac{\nu_e{}^2}{\omega_e{}^2}\right) l_1{}^2} \frac{E^\circ}{B^\circ} \tag{60}$$

The dependence of V on the direction of the wavenormal is evidently sufficiently complicated to make the discussion of the behaviour of a general irregularity

rather elaborate. Attention is here confined to two special cases for which simple results are readily deduced.

As already remarked, there is, of course, no difficulty in the case of an arbitrary one-dimensional irregularity; for V is independent of the wavelength, and hence all the waves forming the irregularity at $t = 0$ travel with the same velocity, which is therefore the velocity of the irregularity as a whole. But this situation is hardly of physical interest.

The one other case giving an equally simple result is that in which the irregularity is two-dimensional and parallel to $\mathbf{B}^\circ$. Though still highly idealized, this

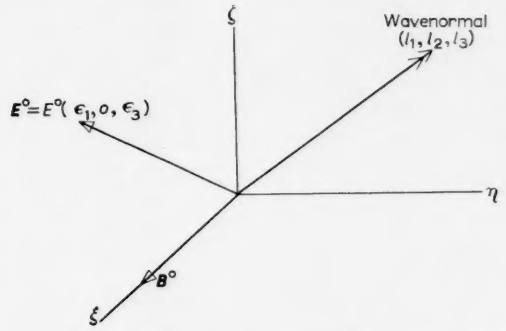


Fig. 2, The configuration to which formula (60) refers.

situation is worth some attention in the light of the remarks at the end of Section 7. It implies that $l_1 = 0$ in equation (60), so:

$$V = \frac{\cos\theta}{1 + \nu_i\nu_e/\omega_i\omega_e}\frac{E^\circ_\perp}{B^\circ} \tag{61}$$

where $E^\circ_\perp$ is the component of $\mathbf{E}^\circ$ perpendicular to $\mathbf{B}^\circ$ and θ is the angle between the wavenormal and $\mathbf{E}^\circ \wedge \mathbf{B}^\circ$ (see Fig. 3).

A more illuminating way of expressing equation (61), is to say that the harmonic wave travels along the direction of $\mathbf{E}^\circ \wedge \mathbf{B}^\circ$ with speed $[1 + \nu_i\nu_e/(\omega_i\omega_e)]^{-1}E^\circ_\perp/B^\circ$; that is, it has velocity:

$$\frac{1}{1 + \nu_i\nu_e/\omega_i\omega_e}\frac{\mathbf{E}^\circ \wedge \mathbf{B}^\circ}{B^{\circ 2}} \tag{62}$$

Since equation (62) is independent both of the wavelength and of the direction of the wavenormal in the plane perpendicular to $\mathbf{B}^\circ$, it gives the velocity of *any* weak two-dimensional irregularity which is parallel to $\mathbf{B}^\circ$.

Because of the rapid variation of the collision frequencies with height, the velocities given by equation (62) are negligible below about 90 km, and effectively equal to $\mathbf{E}^\circ \wedge \mathbf{B}^\circ/B^{\circ 2}$ above about 120 km. A transition thus takes place in the E-region. Below the E-region the wind in the neutral air is the primary motivating

force. Above the E-region a wind is ineffective; for its velocity, $\mathbf{u}$ say, is annulled by that associated with the contribution of $\mathbf{u} \wedge \mathbf{B}^\circ$ to $\mathbf{E}^\circ$.

APPENDIX

Numerical values are given here merely to support the contentions as regards orders of magnitude made in the body of the text. In no sense do they represent an attempt to suggest best estimates.

The values $B^\circ = 4 \times 10^5$ Wb/m^2, $m_i = 4 \times 10^{-26}$ kg, $m_e = 9 \times 10^{-31}$ kg are adopted, from which it follows that $\omega_e = 7 \times 10^6$ sec^{-1}, $\omega_i = 1{\cdot}6 \times 10^2$ sec^{-1},

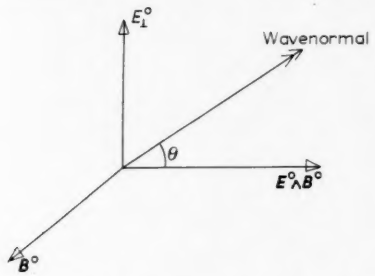


Fig. 3. The configuration to which formula (61) refers.

It is also assumed, for convenience and without significant error, that at all relevant heights $\nu_e = 50\,\nu_i$; this implies $d_e = 890\,d_i$.

Table 1 gives the temperature T, collision frequency ν_e, ionization density N, and mobilities d_e, t_i, t_e at various heights in the ionosphere. The values of N are, of course, only very rough averages.

Table 1.

Height (km)	T (°K)	ν_e (sec^{-1})	N (m)$^{-3}$	d_e (sec-kg^{-1})	t_i (sec-kg^{-1})	t_e (sec-kg^{-1})
70	228	$5{\cdot}2 \times 10^7$	10^8	$2{\cdot}1 \times 10^{22}$	$2{\cdot}4 \times 10^{19}$	$2{\cdot}0 \times 10^{22}$
90	212	$1{\cdot}8 \times 10^6$	10^9	$6{\cdot}1 \times 10^{23}$	$6{\cdot}8 \times 10^{20}$	$3{\cdot}8 \times 10^{22}$
110	240	$8{\cdot}0 \times 10^4$	10^{11}	$1{\cdot}4 \times 10^{25}$	$1{\cdot}6 \times 10^{22}$	$1{\cdot}8 \times 10^{21}$
130	318	$1{\cdot}0 \times 10^4$	10^{11}	$1{\cdot}1 \times 10^{26}$	$7{\cdot}6 \times 10^{22}$	$2{\cdot}2 \times 10^{20}$
150	418	$2{\cdot}0 \times 10^3$	10^{12}	$5{\cdot}6 \times 10^{26}$	$3{\cdot}7 \times 10^{22}$	$4{\cdot}4 \times 10^{19}$
200	700	$1{\cdot}5 \times 10^2$	10^{12}	$7{\cdot}3 \times 10^{27}$	$2{\cdot}9 \times 10^{21}$	$3{\cdot}3 \times 10^{18}$
300	1200	$1{\cdot}2 \times 10$	10^{12}	$9{\cdot}2 \times 10^{28}$	$2{\cdot}3 \times 10^{20}$	$2{\cdot}6 \times 10^{17}$

The order of magnitude of the values of τ (expression (37)) and D (expression (57)) can be obtained by considering the two cases in which the wavenormal is respectively along and perpendicular to $\mathbf{B}^\circ$. In the former, $b_1 = 1$, $b_3 = 0$, and $\tau = \tau_\parallel$, $D = D_\parallel$, say, where

$$\tau_\parallel = \frac{\varepsilon}{Ne^2(d_i + d_e)} \doteq \frac{\varepsilon}{Ne^2 d_e}$$

$$D_\parallel = \frac{2kTd_i d_e}{d_i + d_e} \doteq 2kTd_i$$

VOL. 16 1959

In the latter, $b_1 = 0$, $b_3 = 1$, and $\tau = \tau_\perp$, $D = D_\perp$, say where:

$$\tau_\perp = \frac{\varepsilon}{Ne^2(t_i + t_e)}$$

$$D_\perp = \frac{2kTt_it_e}{t_i + t_e}$$

The values of these expressions at various heights in the ionosphere are given in Table 2. For other directions of the wavenormal, τ lies between $\tau_\parallel$ and $\tau_\perp$, and D lies between $D_\parallel$ and $D_\perp$.

Table 2.

Height (km)	$\tau_\parallel$ (sec)	$\tau_\perp$ (sec)	$D_\parallel$ (m^2-sec^{-1})	$D_\perp$ (m^2-sec^{-1})
70	$1{\cdot}6 \times 10^{-4}$	$1{\cdot}6 \times 10^{-4}$	$1{\cdot}5 \times 10^{-1}$	$1{\cdot}5 \times 10^{-1}$
90	$5{\cdot}6 \times 10^{-7}$	$8{\cdot}0 \times 10^{-6}$	4·0	3·9
110	$2{\cdot}5 \times 10^{-10}$	$1{\cdot}7 \times 10^{-7}$	$1{\cdot}1 \times 10^{2}$	$1{\cdot}1 \times 10$
130	$3{\cdot}1 \times 10^{-11}$	$4{\cdot}1 \times 10^{-8}$	$1{\cdot}1 \times 10^{3}$	1·9
150	$6{\cdot}2 \times 10^{-13}$	$8{\cdot}4 \times 10^{-9}$	$7{\cdot}3 \times 10^{3}$	$5{\cdot}3 \times 10^{-1}$
200	$4{\cdot}7 \times 10^{-14}$	$1{\cdot}1 \times 10^{-7}$	$1{\cdot}6 \times 10^{5}$	$6{\cdot}3 \times 10^{-2}$
300	$3{\cdot}7 \times 10^{-15}$	$1{\cdot}4 \times 10^{-6}$	$3{\cdot}4 \times 10^{6}$	$8{\cdot}6 \times 10^{-3}$

To estimate the upper limit of the order of magnitude of $L = \tau(v_{ix}^\circ - v_{ex}^\circ)$ it can be taken that $|v_{ix}^\circ - v_{ex}^\circ|$ is never greater than 500 m-sec^{-1}. Therefore Lk is everywhere less than 10^{-2} for wavelengths down to 50 m.

Again, from equation (39), the expressions for δ when the wavenormal is respectively parallel and perpendicular to $\mathbf{B}^\circ$ are:

$$\delta_\parallel = \frac{d_i - d_e}{d_i + d_e}, \quad \delta_\perp = \frac{t_i - t_e}{t_i + t_e}$$

Thus $1 - \delta_\parallel^2$ is about $4{\cdot}5 \times 10^{-3}$ at all heights, and $1 - \delta_\perp^2$ is of the same order except in a small height range near 100 km where it sweeps through the value unity. It follows that at all heights above about 75 km D is substantially greater than $L^2(1 - \delta^2)/(4\tau)$, so that diffusive decay is more rapid than that arising from the second term in expression (42) for ω.

No general quantitative statement can be made about the effect of diffusion, since the decay time due to it varies as the square of the wavelength of the disturbance. An impression of its relative importance is gained from Table 3 which gives the wavelength λ for which the decay time associated with diffusion along the magnetic field, $4\pi^2\lambda^2/D_\parallel$, is equal to 1 sec.

Table 3.

Height (km)	70	90	110	130	150	200	300
λ (m)	2·5	$1{\cdot}3 \times 10$	$6{\cdot}6 \times 10$	$2{\cdot}1 \times 10^{2}$	$5{\cdot}4 \times 10^{2}$	$2{\cdot}5 \times 10^{3}$	$1{\cdot}2 \times 10^{4}$

Acknowledgements—The authors would like to express their thanks to Dr. K. WEEKES, Dr. J. W. DUNGEY and Mr. J. P. DOUGHERTY for much helpful discussion.

REFERENCES

CLEMMOW P. C., JOHNSON M. A. and WEEKES K.	1955	*Report of the 1954 Physical Society Conference on the Physics of the Ionosphere*. Physical Society, London.
COWLING T. G.	1957	*Magnetohydrodynamics*. Interscience, New York.
SPENCER M.	1955	*Proc. Phys. Soc. Lond.* **68**, 493.

VOL.
16
1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 37 to 45. Pergamon Press Ltd. Printed in Northern Ireland

V.L.F. reflections from the ionosphere in the presence of a transverse magnetic field

N. F. Barber and D. D. Crombie
Dominion Physical Laboratory,
Department of Scientific and Industrial Research,
Lower Hutt, New Zealand

(*Received* 7 *October* 1958)

Abstract—The reflection coefficient is found for waves that are incident on a sharply bounded ionosphere and have their electric vector in the plane of incidence. The earth's magnetic field is assumed to be horizontal and perpendicular to the direction of propagation. Hence the analysis is appropriate for propagation round the magnetic equator. It is found that the reflection coefficient for waves incident from the west is numerically greater than that for waves incident from the east, when the angle of incidence is large.

Introduction

Crombie (1958) has shown that there is considerable evidence supporting the suggestion by Round *et al.* (1925) that v.l.f. signals propagated over great distances are less attenuated if they travel from west to east than if they travel in the opposite direction. Crombie suggests that this is due to an interaction in the ionosphere between the horizontal component of the earth's magnetic field transverse to the direction of propagation and a component of the electric vector which rotates in the vertical plane of propagation of the wave. This paper calculates the reflection coefficient of a sharply bounded ionosphere at v.l.f. in the presence of a horizontal magnetic field perpendicular to the plane of incidence. The analysis is thus appropriate to propagation around the magnetic equator of the earth.

Budden (1951) and Wait (1957) have determined the reflection coefficient of the ionosphere at v.l.f. when the earth's magnetic field is taken into account. Their results indicate that reversing the direction of propagation does not alter the reflection coefficient of the ionosphere. However, Budden points out that he has ignored the longitudinal component of electric field in the ionosphere. In a later paper, Budden (1954), after calculations on EDSAC resulting from a full wave treatment of the problem, indicates that propagation along the magnetic meridian is reciprocal, but that reciprocity is not to be expected in general.

The magneto–ionic theory of Appleton (1932) discusses a wave whose direction of attenuation happens to be the same as its direction of travel. Fry (1927, 1928) however, has discussed the propagation of optical waves in lossy media (i.e. metals) and shows that the waves penetrating the metal are attenuated in a direction which is different from the direction of phase propagation. Stratton (1941) calls this type of wave an inhomogeneous plane wave. Fry also demonstrates that when the incident wave is polarized in the plane of incidence, the electric vector in the medium rotates in the same plane. Since at v.l.f. the refractive index of the ionosphere is of about the same order of magnitude as the refractive index of metals at optical wavelengths it would be expected that, in the absence

of the earth's magnetic field, the behaviour of v.l.f. waves in the ionosphere would be similar to that of optical waves in metals.

Yabroff (1957) has investigated the reflection of waves at a sharply bounded ionosphere; taking into consideration that the transmitted wave is inhomogeneous. He then assumes that the quasi-longitudinal approximation of Booker (1938) is sufficiently accurate. This prevents the appearance of any non-reciprocal effects.

It will be shown here that when a plane wave, polarized in the plane of incidence, strikes the ionosphere from below, the waves that are excited in the ionosphere are inhomogeneous, and the electrons tend to circulate on elliptic paths lying in the

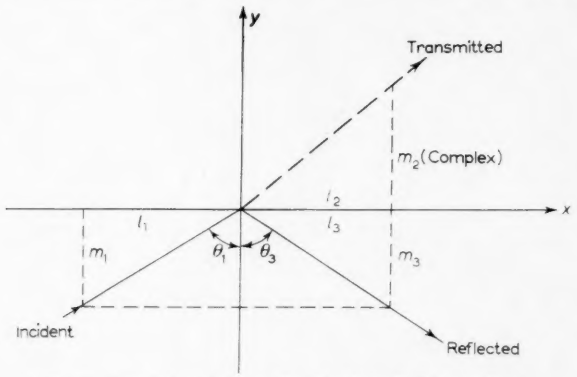


Fig. 1. Reflection from a sharply bounded ionosphere.

plane of incidence. If a horizontal magnetic field is applied, at right angles to the plane of incidence, it affects the motion of the electrons and, depending on its sign, either increases or decreases the area which the electron orbits present to the magnetic field. In general this implies a change in the length of the orbit and hence a change in loss of energy by collisions and a change in the attenuation of the wave in the ionosphere. The electron motion creates a space charge and modifies the electric field of the wave. When the motion is changed by a reversal of the magnetic field, the amplitude and phase of the reflected wave also change. The case treated by Appleton (1932) corresponds to a wave incident normally on the ionosphere, and in this special case the reflection coefficient involves only the second power of the applied magnetic field. Reversing the field does not then affect the reflection.

2. The Wave in the Ionosphere

Consider a plane electromagnetic wave incident from below the sharp horizontal lower boundary of a uniform ionosphere. The incident wave is polarized with its electric vector in the plane of incidence. The incident and reflected waves below the boundary and the transmitted wave above the boundary can all be supposed to vary with space and time in a manner

$$\exp i(pt - lx - my - nz) \tag{1}$$

If the axes of x and z are horizontal then the coefficients p, l, n, must be the same for all waves in order that the boundary conditions may be satisfied everywhere

VOL. 16 1959

at the boundary. By choosing the x,y-plane as the plane of incidence the coefficient n may be set zero for all the waves. Maxwell's equations are

$$\frac{1}{c}\frac{\partial}{\partial t}(\mathbf{E}+\mathbf{P}) = \operatorname{curl}\mathbf{H} \tag{2}$$

$$-\frac{1}{c}\frac{\partial}{\partial t}\mathbf{H} = \operatorname{curl}\mathbf{E} \tag{3}$$

where $\mathbf{E}$ and $\mathbf{H}$ are the electric and magnetic vectors of the wave, and c is the velocity of light in free space. $\mathbf{P}$ is the polarization, that is 4π times the total charge which has been carried in the positive direction across unit area by the drift of the electrons. Since the wave has the variations given by equation (1) the operators $\partial/\partial t$, $\partial/\partial x$, $\partial/\partial y$, $\partial/\partial z$ may be replaced by ip, $-il$, $-im$, $-in$. It has been seen that n can be taken as zero. Equations (2) and (3) then become

$$\left.\begin{aligned} p(E_x + P_x) &= -cmH_z \\ p(E_y + P_y) &= clH_z \\ p(E_z + P_z) &= -c(lH_y - mH_x) \end{aligned}\right\} \tag{4}$$

$$\left.\begin{aligned} pH_x &= cmE_z \\ pH_y &= -clE_z \\ pH_z &= c(lE_y - mE_x) \end{aligned}\right\} \tag{5}$$

Between these equations H_x, H_y, H_z may be eliminated to give

$$\left.\begin{aligned} E_x(p^2 - c^2m^2) + c^2lmE_y &= -p^2P_x \\ E_y(p^2 - c^2l^2) + c^2lmE_x &= -p^2P_y \\ E_z[p^2 - c^2(l^2 + m^2)] &= -p^2P_z \end{aligned}\right\} \tag{6}$$

A further relation between the polarization $\mathbf{P}$ and the electric vector $\mathbf{E}$ is got from the equations of motion of the electrons under the influence of the electric field, $\mathbf{E}$ and the externally impressed magnetic field which is assumed to be in the z direction. A derivativion of these relations is given by MITRA (1952). The form used here is that given by APPLETON.

$$\left.\begin{aligned} E_x &= (\alpha + i\beta)P_x + i\gamma P_y \\ E_y &= (\alpha + i\beta)P_y - i\gamma P_x \\ E_z &= (\alpha + i\beta)P_z \end{aligned}\right\} \tag{7}$$

where

$$\alpha = -\frac{mp^2}{4\pi Ne^2}, \quad \beta = \frac{mpv}{4\pi Ne^2}, \quad \gamma = \frac{mp(eZ/mc)}{4\pi Ne^2}$$

m = mass of electron
e = electronic charge (e.s.u.)
N = number of electrons (per cm^3)
p = angular frequency of the wave (rad/sec)
v = collision frequency (per sec.)
Z = impressed magnetic field in z-direction (G).

From equations (6) and (7) the components of **P** can be eliminated, giving the relationship between the components of **E**. Writing δ for $(\alpha + i\beta)$

$$\left.\begin{aligned} E_x[1 - c^2m^2/p^2 + \delta/(\delta^2 - \gamma^2)] + E_y[c^2ml/p - i\gamma/(\delta^2 - \gamma^2)] &= 0 \\ E_y[1 - c^2l^2/p^2 + \delta/(\delta^2 - \gamma^2)] + E_x[c^2ml/p + i\gamma/(\delta^2 - \gamma^2)] &= 0 \\ E_z\delta[1 - c^2(l^2 + m^2)/p^2] &= 0 \end{aligned}\right\} \tag{8}$$

The last of equation (8) shows that E_z need not be zero if

$$1 = c^2(l^2 + m^2)/p^2$$

This condition does not permit E_x or E_y to be finite, so that it applies to a wave polarized with its electric vector parallel to the impressed field. This wave is not affected by the impressed magnetic field and it will not be discussed further.

If E_x and E_y are to be finite, the determinant of the first two equations of (8) must be zero. This condition can be written

$$p^2/c^2 = (l^2 + m^2)/[1 + (1 + \delta)/(\delta + \delta^2 - \gamma^2)] \tag{9}$$

This applies in or below the ionosphere, provided the appropriate values of δ and γ are used. (Below the ionosphere, $\delta = \infty, \gamma = \infty$.) From the first of equations (8)

$$\frac{E_y}{E_x} = -\frac{1 - c^2m^2/p^2 + \delta/(\delta^2 - \gamma^2)}{c^2lm/p^2 - i\gamma/(\delta^2 - \gamma^2)} \tag{10}$$

From the second of equations (8) a further expression for E_y/E_x can be obtained

$$\frac{E_y}{E_x} = -\frac{c^2ml/p^2 + i\gamma/(\delta^2 - \gamma^2)}{1 - c^2l^2/p^2 + \delta/(\delta^2 - \gamma^2)} \tag{11}$$

Writing the ratio (10) as a/b and the ratio (11) as c/d the combination

$$(la + mc)/(lb + md)$$

leads to the following symmetrical expression:

$$\frac{E_y}{E_x} = -\frac{l(\delta + \delta^2 - \gamma^2) + i\gamma m}{m(\delta + \delta^2 - \gamma^2) - i\gamma l} \tag{12}$$

Equations (7) then give

$$\frac{P_y}{P_x} = -\frac{l(1 + \delta) + i\gamma m}{m(1 + \delta) + i\gamma l} \tag{13}$$

If the ambient magnetic field is absent ($\gamma = 0$) the expressions reduce to

$$E_y/E_x = P_y/P_x = -l/m \tag{14}$$

Now, l is real but m (from equation 9) is complex so long as the electrons suffer collisions ($\beta \neq 0$). Thus even in the absence of an ambient magnetic field the wave in the ionosphere is inhomogeneous since the direction of phase gradient, defined by the real parts of l and m differs from the direction of amplitude attenuation, defined by the imaginary parts of l and m. Further, because the ratio l/m is complex in equation (14), it follows that the electron displacements indicated by

VOL. 16 1959

P_x and P_y are not in phase and that the electrons tend to execute elliptic paths in the plane of incidence.

When a magnetic field is applied, the ratio of the electric components is changed (equation 12) and the ratio of the electron displacements is changed (equation 13). Because both the ratios involve terms having to the first power, a change in the sign of the magnetic field will lead to different electron orbits, as stated in the Introduction.

3. The Reflection Coefficient

The boundary conditions reduce to the condition that H_z and E_x are continuous. To distinguish incident, transmitted and reflected waves their components will be given subscripts 1, 2 and 3 respectively.

The condition (9) for the three waves is

$$\begin{aligned} p^2/c^2 &= l_1^2 + m_1^2 = l_3^2 + m_3^2 \\ &= (l_2^2 + m_2^2)/[1 + (1 + \delta)/(\delta + \delta^2 - \gamma^2)] \end{aligned} \tag{15}$$

Since it has been shown that l_1, l_2 and l_3 must be identical, it follows that

$$\begin{aligned} m_3 &= -m_1 \\ m_2^2 &= m_1^2\left(1 + \frac{1+\delta}{\delta + \delta^2 - \gamma^2}\right) + l_1^2 \frac{1+\delta}{\delta + \delta^2 - \gamma^2} \end{aligned} \tag{16}$$

Since the wave transmitted into the ionosphere cannot increase in amplitude with distance from the boundary the value of m_2 must be such that its imaginary part is negative if y increases upwards. Then if E_x is continuous

$$E_{x1} + E_{x3} = E_{x2} \tag{17}$$

or, from the last of equations (5) and equation (15)

$$H_{z1} - H_{z3} = H_{z2} \cdot \frac{p^2}{c^2 m_1 (m_2 - l_2 E_{y2}/E_{x2})} \tag{18}$$

If H_z is continuous, then

$$H_{z1} + H_{z3} = H_{z2} \tag{19}$$

The reflection coefficient, R, defined as $R = H_{z3}/H_{z1}$ is then obtained from (12) (15), (18) and (19) as

$$R = \frac{m_1[(1+\delta)^2 - \gamma^2] - m_2(\delta + \delta^2 - \gamma^2) + i\gamma l}{m_1[(1+\delta)^2 - \gamma^2] + m_2(\delta + \delta^2 - \gamma^2) - i\gamma l} \tag{20}$$

It will be seen that this expression for R contains both γ and l to the first power. Thus, if the direction of propagation, or the direction of the transverse magnetic field is reversed, the value of R will alter and the reflection is non-reciprocal.

The coefficients m_1 and l are related to the angle θ_1 in Fig. 1 by the relation

$$m_1 : l = \cos\theta_1 : \sin\theta_1$$

Thus m_1, l can be replaced by $p/c \cdot \cos\theta_1$ and $p/c \cdot \sin\theta_1$ respectively. In Fig. 1, if the negative x-direction is west, the positive x-direction is east, and the earth's magnetic field is in the $-z$-direction. Thus γ is negative for W–E propagation.

Moreover, from the form of the waves given in equation (1) it will be seen that for W–E propagation cos θ_1 and sin θ_1 are positive.

The reflection coefficient R from equation (20) is independent of the sign of γ when l or m are zero corresponding to grazing and vertical incidence, respectively.

4. Numerical Results

It has been stated that the conditions investigated in the paper are appropriate to propagation around the magnetic equator, since it is only there that the earth's

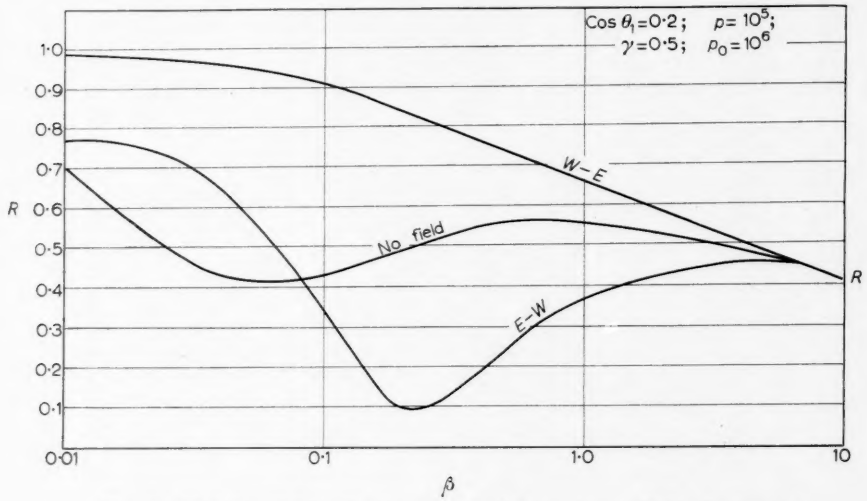


Fig. 2a. Modulus of the reflection coefficient for various collision frequencies.

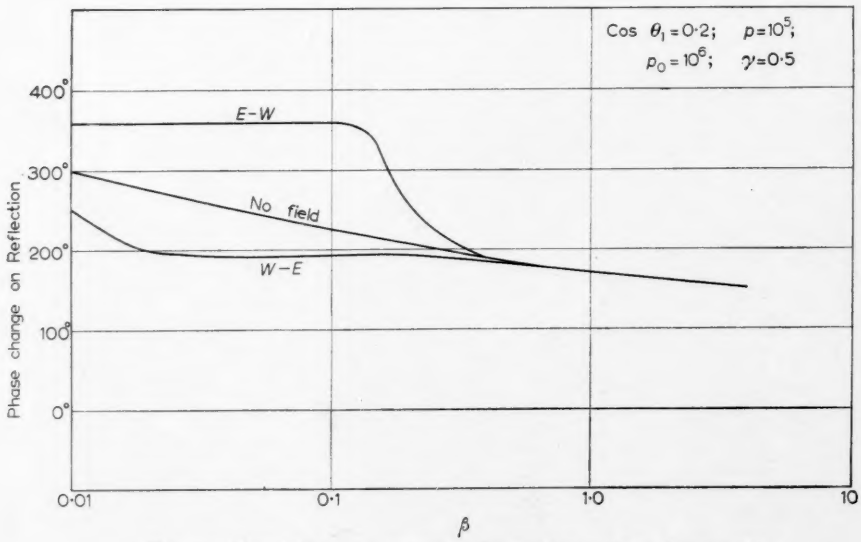


Fig. 2b. Phase change on reflection for various collision frequencies.

VOL. 16 1959

magnetic field is horizontal. Some approximate calculations have been made of the reflection coefficient for W–E and E–W propagation for various conditions. Fig. 2 shows the magnitude and phase of the reflection coefficient for W–E and E–W propagation when the angle of incidence is 78·5°. The wave angular frequency has been taken as 10^5/sec while the gyro angular frequency (eZ/mc) has been taken as 5×10^6/sec. The critical angular frequency of the ionosphere has been assumed to be 10^6/sec and the collision frequency has values between 5×10^5 and 10^8/sec.

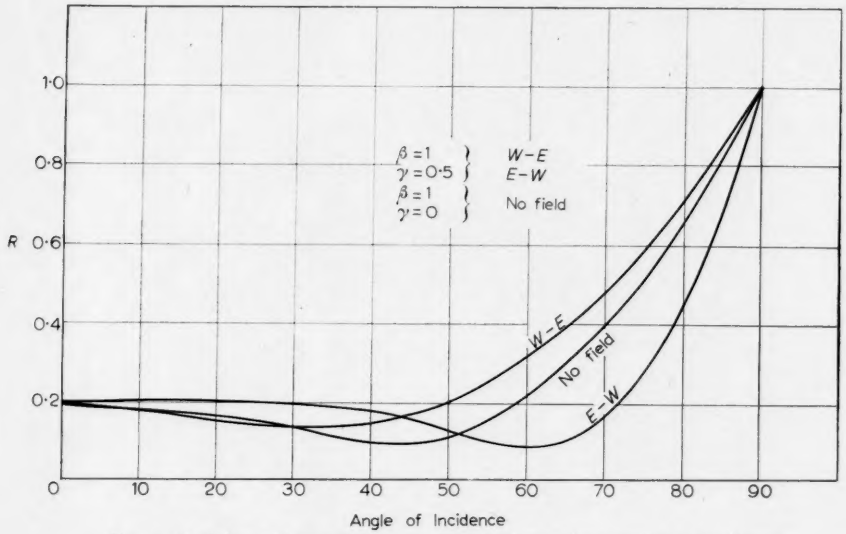


Fig. 3. Modulus of the reflection coefficient for various angles of incidence.

Also plotted for comparison are the magnitude and phase of the reflection coefficient in the absence of the magnetic field. It will be seen that for large values of β, the magnetic field has little effect, but as the collision frequency decreases the E–W and W–E reflection coefficient depart from the no-field value. For values of $\beta = 0{\cdot}2$ the difference in reflection coefficients is pronounced.

It will be observed that there is little difference in the phase change on reflection until β is less than about 0·3. For values of β less than this the change of phase for E–W propagation rapidly increases to 360°, the value for a perfect conductor. On the other hand, the W–E phase change remains nearly equal to 180° until β falls to a very small value.

Fig. 3 shows the effect of varying the angle of incidence while the other parameters are fixed. At large angles the W–E reflection coefficient is greater than either the E–W or no-field reflection coefficient. For small angles of incidence, however, the W–E and no-field reflection coefficients become nearly equal and less than the E–W reflection coefficient.

The magnitude of the W–E and E–W reflection coefficients are approximately equal at an angle of incidence for which the imaginary part of

$$m_1[(1 + \delta)^2 - \gamma^2] - m^2(\delta + \delta^2 - \gamma^2)$$

is zero. Then the magnitude of the numerator of equation (20) is independent of the sign of γl while the magnitude of the denominator does not change appreciably when the sign of γl is changed. For this angle of incidence, however, the phase change on reflection depends on the direction of propagation.

In Fig. 4 the effect of varying the frequency of the incident wave is shown, the electron density, collision frequency, and transverse magnetic field remaining fixed. The ratio of the W–E and E–W reflection coefficients rises slowly from 4 kc/s and has a broad maximum about 16 kc/s falling more rapidly at 32 kc/s.

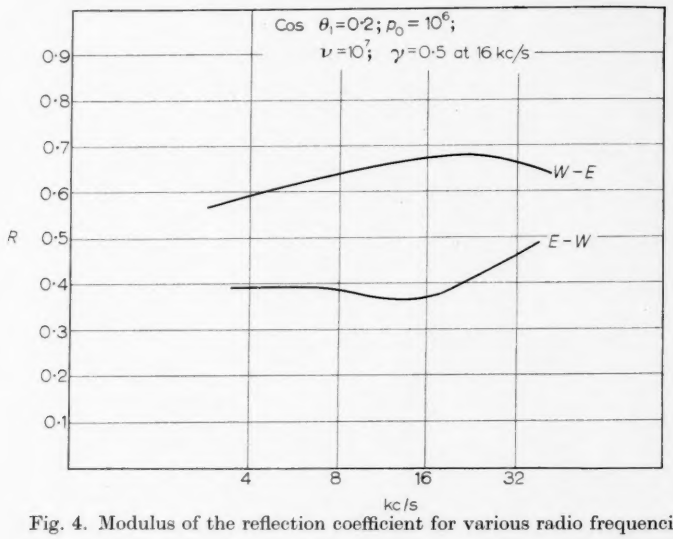


Fig. 4. Modulus of the reflection coefficient for various radio frequencies.

The change of phase on reflection for both W–E and E–W reflection is substantially constant. These results are not entirely in agreement with the experimental data (Crombie, 1958) which suggested that the non-reciprocity was small at frequencies above 20 kc/s and became larger with decreasing frequency down to 12 kc/s. The observations refer, however, to paths that are far from being equatorial.

Discussion

It will be seen from equation (9) that the dielectric constant of the ionosphere is

$$\varepsilon = 1 + 1/[\delta - \gamma^2/(1 + \delta)]$$

which is identical with the expression obtained by Appleton when the magnetic field is transverse. This expression is independent of the sign of γ. It is perhaps surprising that reflection from a medium having such a dielectric constant does depend on the sign of γ. This question has been discussed by Suhl and Walker (1954) who show that certain geometrical symmetry is required, in the presence of a transverse magnetic field, in order that reciprocity be preserved. Consider for example, multiple reflections between two parallel slabs of an ionized gas in the presence of a transverse magnetic field. If the wave experiences an equal number of reflections at each surface, then the amplitude of the emerging wave will not

VOL. 16 1959

alter if the direction of propagation is reversed. This is a highly symmetrical arrangement. If, however, one of the ionized slabs is replaced by the earth's surface, the symmetry is removed and the wave attenuation will depend on the direction of propagation.

The present work is an investigation of the very special care of propagation along the magnetic equator. For this reason there are no experimental data which can be used for direct comparison with the results deduced in this paper. It is believed, however, that there is a v.l.f. transmitter (NBA) at Balboa, which is near the magnetic equator. This transmitter would be a useful source of equatorial v.l.f. propagation studies, but it does not appear to be active at present.

Alternatively, it might be possible to make observations at the magnetic equator, of signals radiated from lightning. It has been suggested that the directional asymmetry of atmospherics (HEPBURN, 1957) might be due to non-reciprocity of the type discussed in this paper. Observations of atmospheric waveforms using directional aerials near the magnetic equator would establish whether this is true more easily than the same observations elsewhere, since the theory of propagation round the equator seems much simpler than for other directions.

CONCLUSION

It has been shown that in circumstances appropriate to propagation at very low frequencies, round the magnetic equator, the reflection coefficient of the ionosphere depends on whether the direction of propagation is from E–W or from W–E. This provides some theoretical support for the experimental observation reported earlier, although these observations refer to oblique rather than equatorial paths. There is no experimental evidence which is directly applicable to the conditions considered in the paper. Possible ways in which direct evidence could be obtained are outlined.

REFERENCES

APPLETON E. V.	1932	*J. Instn. Elect. Engrs.* **71,** 642.
BOOKER H. G.	1938	*Proc. Roy. Soc.* A **150,** 267.
BUDDEN K. G.	1951	*Phil. Mag.* **42,** 833.
BUDDEN K. G.	1954	*Proc. Camb. Phil. Soc.* **50,** 604.
CROMBIE D. D.	1958	*J. Atmosph. Terr. Phys.* In press.
FRY T. C.	1927	*J. Opt. Soc. Amer.* **15,** 137.
FRY T. C.	1928	*J. Opt. Soc. Amer.* **16,** 1.
HEPBURN F.	1957	*J. Atmosph. Terr. Phys.* **10,** 1.
MITRA S. K.	1952	*The Upper Atmosphere* (2nd Ed.). The Asiatic Society, Calcutta.
ROUND H. J., ECKERSLEY T. L., TREMELLEN K. G. and LUNNON F. C.	1925	*J. Instn. Elect. Engrs.* **63,** 933.
STRATTON J. A.	1941	*Electromagnetic Theory.* McGraw-Hill, New York.
SUHL H. and WALKER L. R.	1954	*Bell Syst. Tech. J.* **33,** 939.
WAIT J. R. and PERRY L. B.	1957	*J. Geophys. Res.* **62,** 43.
YABROFF I. W.	1957	*Reflection and Transmission at a Sharply-Bounded Ionosphere.* Radio Propagation Laboratory, Stanford University, U.S.A.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 46 to 58. Pergamon Press Ltd. Printed in Northern Ireland

A discussion of the motion in nitrogen of free electrons with small energies with reference to the ionosphere

L. G. H. HUXLEY
University of Adelaide

(*Received* 15 *October* 1956)

Abstract—In this paper measurements of diffusion and drift of slow electrons in pure nitrogen carried out by CROMPTON and HALL are discussed. The discussion is restricted to the range of energies of the electrons between that of thermal equilibrium and four times that value. From the behaviour of the coefficient of diffusion D it is concluded that in elastic collisions between electrons and nitrogen molecules the collisional cross-section A is proportional to the speed c of the electron; namely, $A = 3{\cdot}29 \times 10^{-23} c$ cm^2. From this result an accurate expression is derivable for the velocity of drift W in terms of the parameter Z/p and the mean energy of agitation of the electrons. The measured rate w at which electrons with energy Q lose energy in collisions with molecules is accurately described by a formula that ascribes these losses of energy chiefly to the excitation of changes in the rotational states of the molecules.

A formula in close agreement with the measurements is derived for the dependence of the mean energy of agitation of the electrons upon Z/p.

The significance of the measurements for ionospheric studies is considered with particular reference to the interpretation of measurements of radio-wave interaction. It is shown that the values of the molecular concentration in the height range 82 to 90 km above the ground, found by means of radio-wave interaction, are consistent with the A.R.D.C. model atmosphere (MINZNER and RIPLEY, 1956).

VOL. 16 1959

1. AGITATIONAL ENERGY AND COEFFICIENT OF DIFFUSION

THE motions of electrons in pure nitrogen have recently been investigated experimentally in this laboratory by CROMPTON and HALL and in what follows some aspects of their measurements* are discussed in detail.

The experimental procedure was a development of that introduced by TOWNSEND in which the ratio W/D of the drift speed W of the electrons to the coefficient of diffusion D is directly measured over a range of electric field strengths Z and pressures p of the gas.

In addition, separate measurements were made in an electrical shutter apparatus, of the drift speed W as a function of the ratio Z/p. It is evident that for any combination of field strength Z and pressure p, the coefficient of diffusion D may be separately deduced as the product of the corresponding measured values of W and D/W.

Formulae for W and D in terms of the mean free path $l(c)$ of the electrons or the collisional cross-sections $A(c)$ of the molecules, and the agitational speeds c of the electrons are (e.g. HUXLEY, 1957a):

$$\left.\begin{aligned} W &= (Ze/3m)\overline{[c^{-2}d/dc(lc^2)]} = (Ze/3Nm)\overline{[c^{-2}d/dc(c^2/A)]} \\ \text{and} \qquad D &= (\overline{lc})/3 = (\overline{c/A})/3N \end{aligned}\right\} \qquad (1)$$

in which, e/m is the specific charge of the electron, and N is the number of molecules in unit volume of the gas. The average, denoted by the superimposed bar, is taken

* To be published.

with respect to the distribution of the speeds c. From equations (1) it follows that:

$$W/D = [3Ze/2Q]h \tag{2}$$

where $Q = \overline{mc^2}/2$ is the mean energy of agitation of an electron and the factor $h = [c^2 \,.\, \overline{c^{-2}d/dc(lc^2)}]/[3\overline{(lc)}]$.

The dimensionless factor h is determined both by the law of distribution of the speeds c and by the dependence of the mean free path $l \equiv l(c)$, upon c.

It can be shown (HUXLEY, 1957a) that if the law of distribution of the speeds c is that of Maxwell, then the factor h is equal to unity, in which event:

$$W/D = 3Ze/2Q$$

so that the value of Q for a group or stream of electrons moving through the gas in a steady state of motion in a known field Z may be found from the measured value of W/D. In general the law of distribution of the speeds is not that of Maxwell and h is not unity.

What is determined therefore from measurements of W/D in known fields Z is the apparent kinetic energy Q_a defined as:

$$Q_a = Q/h \tag{3}$$

that is, in order to find the correct value of $Q = hQ_a$ from the measured quantity Q_a it is necessary to know the law of distribution of the speeds c and the dependence $l(c)$ of l upon c in order to determine h.

Consider next the coefficient of diffusion D as represented by the second of equations (1). If p is measured in millimetres of mercury and the laboratory temperature is 288°K, then the molecular concentration N may be expressed as $N = 3{\cdot}35 \times 10^{16}p$; consequently the formula for D may be transformed to:

$$\left.\begin{aligned} pD &= \overline{(c/A)}/(1{\cdot}00 \times 10^{17}) = \overline{(c/A)} \times 10^{-17} \\ \text{Similarly} \qquad W &= (e/m)(1/3{\cdot}35 \times 10^{16})(Z/p)[c^{-2}d/dc(c^2/A)] \end{aligned}\right\} \tag{4}$$

Experiment shows that W is accurately a function of the ratio Z/p, from which it follows that pD, h, Q, Q_a and the law of distribution of the speeds c are determined, in a steady state of motion by the value of Z/p, that is to say, by Z/N. Thus Z/p or Z/N are convenient parameters through which W, pD, Q_a, etc., may be related to each other. It is sometimes convenient to express the values of Q and Q_a in terms of the mean thermal energy of agitation Q_0 of the molecules as unit. When measured in this unit Q and Q_a will be denoted by $q = Q/Q_0$; $q_a = Q_a/Q_0$. At a temperature of 288°K, $Q_0 = 5{\cdot}96 \times 10^{-14}$ ergs which is approximately $\frac{1}{27}$ eV.

2. Experimental Results in Nitrogen when $Z/p < 0{\cdot}1$

The values of two of the quantities discussed above at the smaller values of Z/p as measured by CROMPTON and HALL are shown in Table 1 (Z in V/cm; p in mm Hg).

It is evident that when $q_a < 4$, pD is remarkable in its constancy. The measured values show that in the range of q_a under consideration the diffusion coefficient in nitrogen is closely represented by:

$$pD = 3{\cdot}02 \times 10^5 \tag{5}$$

(D in cm^2 sec^{-1})

It follows therefore that in nitrogen $\overline{(c/A)} = 3{\cdot}04 \times 10^{22}$. The situation is that the mean value of a function of c is a constant independent of the mean value $\bar{c}$ of c; it follows that the ratio c/A is itself independent of c and is a constant equal to

Table 1

Z/p	0·02	0·03	0·04	0·05	0·06	0·07	0·08	0·09
$q_a = Q_a/Q_0$	1·29	1·56	1·87	2·20	2·56	2·90	3·26	3·65
$pD \times 10^{-5}$	3·04	2·99	3·04	2·98	3·01	3·02	3·05	3·10

$3{\cdot}04 \times 10^{22}$. The collisional cross-section A in nitrogen is therefore proportional to c and is given by the formula:

$$[q_a < 4; \quad T = 288°\text{K}]; \quad A = 3{\cdot}29 \times 10^{-23} c \text{ cm}^2 \tag{6}$$

The proportionality of A to c when c is small, was noticed in earlier measurements (Crompton, *et al.* 1953).

VOL. 16 1959

3. Interpolation Formula for W

When two quantities, such as W/D and W, are found both by experiment as functions of Z/p by different techniques it is to be expected that in general they will not be measured with equal precision, consequently it is a fortunate circumstance when it proves possible, for physical reasons, to express the less accurate quantity in terms of the more accurate.

In the present instance the physical fact of the proportionality of A to c implies the constancy of Dp whose value can be found accurately. Of the quantities W/D and W the former can be measured, when $q < 4$, with greater precision than the latter, and it is therefore useful, using the fact of the constancy of Dp to deduce the following expression for W:

$$W = (W/D)(Dp)/p = 3{\cdot}02 \times 10^5 (W/D)/p \tag{7}$$

from which W may be found as a function of Z/p.

An equivalent form of this equation that is derived by replacing W/D by its value given by equation (2) is the following:

$$W = 1{\cdot}22 \times 10^7 (Z/p)/q_a \text{ cm sec}^{-1} \tag{8}$$

(Z in V/cm).

These formulae may be used to interpolate and extrapolate values of W within and below the range of measured values. The reliability of the formulae

for this purpose can be judged by the way in which the measured values are recovered from it, as shown in Table 2.

Table 2

Z/p	0·02	0·03	0·04	0·05	0·06	0·07	0·08
q_a	1·29	1·56	1·86	2·20	2·56	2·90	3·26
$W_{obs} \times 10^{-5}$	1·9	2·32	2·59	2·73	2·84	2·94	3·02
$W \times 10^{-5}$ (from formula)	1·88	2·35	2·59	2·75	2·84	2·92	2·97

4. Motions of Electrons in Air when Z/p is Small

Collisional cross-section and collisional frequency

If A_{air}, $A_{nitrogen}$ and A_{oxygen} are the collisional cross-sections of molecules of air, nitrogen and oxygen respectively, in collisions with electrons with a speed c, then these cross-sections should be related as follows:

$$A_{air} = 0{\cdot}8A_{nitrogen} + 0{\cdot}2A_{oxygen} \tag{9}$$

A limiting case would be that in which the term $0{\cdot}2\,A_{oxygen}$ was negligibly small so that $A_{air} = 0{\cdot}8A_{nitrogen}$. In that event, according to equation (6), A_{air} would be given by

$$A_{air} = 2{\cdot}63 \times 10^{-23}c \text{ cm}^2 \qquad (q < 4)$$

A linear dependence of A_{air} upon c had been suggested on the basis of measurements of the smallest values of q that had been reached (Crompton *et al.* 1953) and these authors proposed the relationship $A_{air} = 2{\cdot}48 \times 10^{-23}c$ cm^2.

When this value is substituted in the formula $0{\cdot}2A_{oxygen} = A_{air} - 0{\cdot}8A_{nitrogen}$ along with the value of $A_{nitrogen}$ given in equation (6), the value of A_{oxygen} that is deduced is negative.

However, the values of A_{air} were measured at values of $q > 4$, consequently there is no overlap of the experimental range of q in air and that in nitrogen ($q < 4$), within which $A_{nitrogen}$ is given accurately as $3{\cdot}29 \times 10^{-23}c$ cm^2. Moreover a more careful scrutiny of the dependence of A_{air} upon q shows that A/c is not strictly constant in the neighbourhood of $q = 4$, but increases slowly as q diminishes so that, were the trend maintained into the vicinity of $q = 3$, the value of A/c would not be very different from $2{\cdot}63 \times 10^{-23}$, which it must exceed if negative values of A_{oxygen} are not to be deduced.

Although it is not possible to deduce the values of A_{oxygen} ($q < 4$) it may reasonably be concluded that $0{\cdot}2A_{oxygen}$ is small compared with $0{\cdot}8A_{nitrogen}$ and that $A_{air} \doteq 0{\cdot}8A_{nitrogen}$.

It will be assumed that in the absence of direct measurements in oxygen, that this conclusion is correct over the whole range $1 \leq q \leq 4$.

It is, unfortunately, difficult to make trustworthy measurements of W/D and of W in air when $q < 4$ because of the influence of the process of the attachment of electrons to molecules to form negative ions so that there is no immediate

prospect that these quantities will be measured for values of q appreciably less than four.

The collisional cross-sections deduced from D and discussed above refer to the collisions that are quasi-elastic; however, in a small proportion of the total number of collisions in a given interval some electrons lose a relatively large proportion of their energy of agitation when they excite changes in the rotational and other states of the molecules. Although the collisional cross-sections of the molecules of oxygen in elastic collisions are considered not to contribute to the effective cross-sections for air, yet this is not of necessity true for the rate at which energy is lost in all collisions, for collisions made by those few electrons with energies notably greater than the mean energy of agitation can produce changes, when $Q < 4Q_0$, in the rotational states of the oxygen molecules. The mean rate at which energy is lost by an electron in air when $q > 4$ is not to be deduced in a simple manner from the rates in oxygen and nitrogen separately. Reasons are given in Section (7) for believing that when $q < 4$ the rate of loss of energy of an electron in air is due chiefly to collisions with the molecules of nitrogen.

5. The Mean Rate at which Electrons Lose Energy in Nitrogen

In a steady state of motion the mean energy Q of the electrons in a group remains constant, the motion being one of equilibrium in which the mean rate at which electrons acquire energy as they drift in an electric field Z is the same as that at which they lose it in collisions with the molecules of the gas. The rate at which energy is given to an electron in the mean is ZeW so that if w is the mean rate at which it loses energy in collisions then:

$$w = ZeW = ep(Z/p)W$$

that is

$$w/p = e(Z/p)W \tag{10}$$

It follows that w/p is derived immediately as a function of Z/p from the experimental values of W in terms of Z/p.

If use is made of equation (7) then in nitrogen ($q_a < 4$).

$$\begin{aligned} w &= 3{\cdot}02 \times 10^5 e(W/D)(Z/p)/300 \\ &= 4{\cdot}83 \times 10^{13}(W/D)(Z/p) \text{ erg sec}^{-1} \quad (Z \text{ in V-cm}^{-1}) \end{aligned}$$

A relation between w, Z/p and q_a follows from equation (8); it is ($q_a < 4$).

$$w/N = 5{\cdot}82 \times 10^{-22}(Z/p)^2/q_a \tag{11}$$

6. The Mean Energy Lost in a Collision Averaged over all Collisions

The average number of collisions made by an electron in a second is the mean of the collisional frequencies, $\nu = NAc$. But, from equation (6), in nitrogen when $q_a < 4$ ($T = 288°\text{K}$) $A = 3{\cdot}29 \times 10^{-23}c$, consequently:

$$\begin{aligned} \bar{\nu} &= 3{\cdot}29 \times 10^{23} N\overline{c^2} = 7{\cdot}22 \times 10^4 NQ \\ &= 4{\cdot}30 \times 10^{-9} Nq \end{aligned} \tag{12}$$

VOL. 16 1959

The proportion of the mean energy Q lost in a collision averaged over all collisions is:

$$\eta = w/\tilde{\nu}Q$$

But, in nitrogen at $T = 288°\text{K}$, $\tilde{\nu}Q = 2{\cdot}56 \times 10^{-22}Nq^2$ $(q < 4)$, therefore:

$$\eta = 3{\cdot}91 \times 10^{21}(w/N)/q^2 \tag{13}$$

If w/N is given its value (11) then,

$$\eta = 2{\cdot}28(Z/p)^2/q^2q_a \tag{14}$$

When the distribution of the speeds c is Maxwellian then $q^2q_a = q^3$.

7. Theoretical Formulae for w/N

When q_a is small, as in the present discussion, the total mean rate w at which an electron loses energy in collisions is the sum of two contributions, the one w_e representing the mean rate at which energy is lost in elastic collisions and the other w_r, the mean rate in collisions that produce rotational changes in molecules (Huxley, 1956).

The rate of energy loss $\mathrm{w_e}$

If the collisions, in which changes in the internal energy of the molecules do not occur, resemble those between smooth perfectly elastic spheres, then the proportion of its energy lost by an electron of a group (Maxwell's distribution of speeds) in a collision is, on the average:

$$\eta_e = (2{\cdot}66m/M)(1 - 1/q)$$

where m/M is the ratio of the mass of an electron to that of a molecule. It follows that

$$w_e = \nu Q\eta_e = (2{\cdot}66m/M)\tilde{\nu}(Q - Q_0)$$

In nitrogen, when $q < 4$, $\tilde{\nu} = 7{\cdot}22 \times 10^4 NQ$ (equation 12) consequently, after inserting the appropriate value for m/M, it is found that

$$\begin{aligned} w_e/N &= 3{\cdot}78Q(Q - Q_0) \\ &= 3{\cdot}78q(q - 1)Q_0^2 \text{ erg sec}^{-1} \text{ cm}^3 \end{aligned} \tag{15}$$

At $T = 288°\text{K}$ where $Q_0 = 5{\cdot}96 \times 10^{-14}$ erg this becomes:

$$w_e/N = 1{\cdot}33 \times 10^{-26}q(q - 1) \text{ erg sec}^{-1} \text{ cm}^3 \tag{16}$$

The rate of energy loss $\mathrm{w_r}$

Theoretical consideration of the rate w_r led to the following semi-empirical formula for w_r/N (Huxley, 1956), when the distribution of the speeds c follows the law of Maxwell

$$\begin{aligned} w_r/N &= \alpha Q^{\frac{1}{2}}[\exp(-\beta/Q) - \exp(-\beta/Q_0)] \\ &= aq^{\frac{1}{2}}[\exp(-b/q) - \exp(-b)] \end{aligned} \tag{17}$$

in which α and β are constants independent of Q_0 and $a = \alpha/Q_0^{\frac{1}{2}}$ and $b = \beta/Q_0$.

The associated proportional loss in a collision averaged over all collisions $\bar{\nu}$ is therefore

$$\eta_r = w_r/\bar{\nu}Q = dq^{-3/2}[\exp(-b/q) - \exp(-b)] \tag{18}$$

Since in nitrogen, at $T = 288°\text{K}$, $\bar{\nu}Q = 2{\cdot}56 \times 10^{-22}Nq^2$ it follows that $a = 2{\cdot}58 \times 10^{-22}d$ at that temperature.

The constants a and b, or d and b are found from the experimental values of w_r or η_r. In practice it proves to be more convenient to find d and b from η_r determined as a function of q_a ($= q$ since Maxwell's distribution is postulated) than from w_r

Table 3

(The rows to be compared are linked)

Z/p	0·02	0·03	0·04	0·05	0·06	0·07	0·08
q	1·29	1·56	1·87	2·20	2·56	2·90	3·26
w/N from (11); $\times 10^{26}$	18·0	33·6	49·8	66·1	81·8	98·4	114·0
w_e/N from (16); $\times 10^{26}$	0·50	1·16	2·16	3·51	5·31	7·33	9·80
$w_r/N = (w - w_e)/N$; $\times 10^{26}$	17·5	32·4	47·6	62·6	75·5	91·0	104
w_r/N from (17); $\times 10^{26}$	17·5	32·3	47·8	62·6	77·4	90·8	103

measured as a function of q. The curve representing the dependence of η_r upon q shows that η_r possesses a maximum value $(\eta_r)_{\max}$ at $q = q_{\max}$ whereas w_r possesses no maximum. The condition for a maximum value of η_r may be derived from equation (18) in the form of a transcendental equation relating b to $q_{\max}$, consequently b can be found from the experimental value of $q_{\max}$ which in nitrogen is 1·55. This value gives $b = 0{\cdot}74$ for this gas. The value of d is then determined from formula (18) using this value of b and the measured value of $(\eta_r)_{\max}$. It was found that $d = 6{\cdot}90 \times 10^{-3}$; the value of a follows from that of d and is $a = 1{\cdot}78 \times 10^{-24}$. These values, as already explained, refer to a temperature of 288°K.

The values of α and β in equation (17) are deduced from those of a and b by putting $Q_0 = 5{\cdot}96 \times 10^{-14}$ erg. It is found that:

$$\alpha = 7{\cdot}29 \times 10^{-18}; \qquad \beta = 4{\cdot}41 \times 10^{-14}$$

Although these constants of the formulae (17) and (18) were found from the co-ordinates of a single point $((\eta_r)_{\max}, q_{\max})$ on the $(\eta_r q)$ curve yet the formulae agree closely with the measured values of w_r and η_r over a wide range, a fact which gives confidence in these formula when they are used for interpolation or extrapolation to small values of q.

The reliability of the formulae may be judged from the comparisons between measured and calculated values made in Table 3.

VOL. 16 1959

The total rate of loss of energy in collisions $w = w_r + w_e$ is (in erg sec^{-1} cm^3) therefore given accurately by ($q < 4$);

$$\begin{aligned} w/N &= 7{\cdot}29 \times 10^{-18} Q^{\frac{1}{2}}[\exp(-4{\cdot}41 \times 10^{-14}/Q) - \\ &\quad - \exp(-4{\cdot}41 \times 10^{-14}/Q_0)] + 3{\cdot}78Q(Q - Q_0) \\ a &= 1{\cdot}78 \times 10^{-24} \\ b &= 0{\cdot}74 \end{aligned} \tag{19}$$

and at $T = 288°$K by:

$$w/N = 1{\cdot}78 \times 10^{-24} q^{\frac{1}{2}}[\exp(-0{\cdot}74/q) - \exp(0{\cdot}74)] + 1{\cdot}33 \times 10^{-26} q(q-1) \tag{20}$$

Approximate forms when q $\rightarrow$ 1

When $(q-1)/q \ll 1$, equations (19) and (20) become:

$$w/N = [\alpha\beta/Q_0^{3/2}) \exp(-\beta/Q_0)](Q_0/Q)^{\frac{1}{2}}(Q - Q_0) + 3{\cdot}78Q(Q - Q_0) \tag{21}$$

with $\alpha = 7{\cdot}29 \times 10^{18}$ and $\beta = 4{\cdot}41 \times 10^{-14}$ as before;
also, at $T = 288°$K

$$\begin{aligned} w/N &\rightarrow [ab \exp(-b)]q^{\frac{1}{2}}(q-1) + 1{\cdot}33 \times 10^{-26} q(q-1) \\ &= (6{\cdot}28 \times 10^{-25} q^{-\frac{1}{2}} + 1{\cdot}33 \times 10^{-26} q)(q-1) \end{aligned} \tag{22}$$

The factors Q/Q_0 and q in equations (21) and (22) are approximately equal to unity.

The smallest value of Z/p, shown in Table 3, at which q is 1·29, is that at which the smallest measured value of the drift velocity W was made; consequently it was not possible to derive the quantity w/N directly from the measurements for smallest measured values q, although these were obtained down to $q_a = 1{\cdot}03$. However, by making use of the result $A = 3{\cdot}29 \times 10^{-23} c$ (equation 6) values of w/N were derived from equation (11) from the measured values of q_a as a function of Z/p. These values of w/N, right down to $q_a = 1{\cdot}03$ were in close agreement with the values predicted by formula (20).

8. Formula Relating q_a to (Z/p) in Nitrogen at $T = 288°$K

When w/N is eliminated from formulae (11) and (22) the following formula is obtained:

$$\begin{aligned} (Z/p)^2 = 3{\cdot}06 \times 10^{-3} q^{3/2}[\exp(-0{\cdot}74/q) - \exp(-0{\cdot}74)] + \\ + 2{\cdot}29 \times 10^{-5} q^2(q-1) \end{aligned} \tag{23}$$

The experimental values of q and Z/p in nitrogen are recovered from this formula with a high degree of accuracy within the range $1{\cdot}03 \leq q \leq 4$. When $q \rightarrow 1$ the formula becomes

$$(Z/p)^2 = 1{\cdot}065 \times 10^{-3} q^{\frac{1}{2}}(q-1) + 2{\cdot}29 \times 10^{-5} q^2(q-1) \tag{24}$$

9. Applications to Ionospheric Studies

A knowledge of the features of the motions of free electrons in air when q does not appreciably exceed the value unity, corresponding to the state of thermal

equilibrium, would be significant for studies of ionospheric absorption and radio-wave interaction. Unfortunately the process of electron attachment makes the study of electron motion in air when $q < 4$ difficult and imprecise and it would seem that the most promising procedure at present is to deduce the nature of the motion of electrons in air when $q < 4$ from what is known of their motions in pure nitrogen and in air and oxygen. This procedure is adopted in what follows.

Collisional frequency of electrons in air

In Section (2) it was shown that in nitrogen the collisional cross-section A of a molecule was proportional to the speed c of the colliding electron; specifically:

$$A = 3{\cdot}29 \times 10^{-23}c \qquad (q < 4;\ T = 288°\text{K})$$

from which it was deduced (equation 12) that the mean collisional frequency $\bar{\nu}$ is given by

$$\bar{\nu} = 7{\cdot}22 \times 10^{4}NQ$$

From the behaviour of electrons in air it was inferred in Section (6) that the collisional cross-sections in elastic collisions of the molecules of oxygen, when $q < 4$, are relatively small compared with those of nitrogen in the majority of collisions at the same speeds c, and that for the purpose of estimating collisional frequencies (when $q < 4$) air may be regarded as nitrogen diluted by the factor 0·8. The collisional frequency in air is therefore given by:

$$\bar{\nu} = 0{\cdot}8 \times 7{\cdot}22 \times 10^{4}NQ = 5{\cdot}8 \times 10^{4}NQ$$

In a state of thermal equilibrium Q becomes Q_0. However, the physical pressure p of the gas is $p = (2/3)NQ$ consequently the collisional frequency of thermal electrons is given by:

$$\left.\begin{aligned} &\bar{\nu} = 8{\cdot}7 \times 10^{4}p \qquad (p \text{ is in dyn cm}^{-2}) \\ &\text{When } p \text{ is expressed in mm Hg the formula is} \\ &\bar{\nu} = 1{\cdot}2 \times 10^{8}p \end{aligned}\right\} \qquad (25)$$

For instance in a region of the ionosphere where $p = 10^{-3}$ mm Hg the collisional frequency $\bar{\nu}$ of thermal electrons is $\bar{\nu} = 1{\cdot}2 \times 10^{5}\ \text{sec}^{-1}$.

The coefficient $1{\cdot}2 \times 10^{8}$ in equation (25) differs from that previously given by Crompton *et al.* (1953), which was $8{\cdot}4 \times 10^{7}$. The discrepancy is due chiefly to the of an unsatisfactory method of averaging that had been adopted.

Rate of loss of energy by electrons in air

Behaviour of η in oxygen. When the experimental dependence of η upon q for oxygen is displayed graphically it is found that η rises rapidly to a maximum at value of q of approximately 13. On the other hand since the moments of inertia of the molecules of oxygen and nitrogen have comparable values it is to be expected that the maximum value of η_r would be found in oxygen as in nitrogen, in the vicinity of $q = 1{\cdot}8$. The measurements in oxygen have not yet been made for values of q less then 4·7, consequently it is not possible to assert that such a maximum value of η_r exists. Despite the fact that a formula of the same type as

VOL. 16 1959

equation (20) fitted the measurements it would seem nevertheless that the maximum value of η at $q = 13$ is due to the excitation of changes in the vibrational state of the oxygen molecule rather than in exciting rotational changes as was at first supposed (HUXLEY, 1956). Since the reduction in the value of η in oxygen is very rapid as q is diminished it will be assumed that when q is less than 4 the losses represented by η are not vibrational losses but true rotational losses η_r, and, moreover, that these give values of η_r as functions of q that are nearly the same as those in nitrogen; that is to say unlike the elastic cross-sections, the collisional cross-sections for rotational changes are comparable in oxygen and nitrogen for a given value of q.

Table 4

q		4·7	5	6	10
$(w/N)_{\text{oxygen}}$		7·74	9·0	15·2	59
$(w/N)_{\text{nitrogen}}$	$\times 10^{26}$	1·98	2·1	2·8	6·2
$(w/N)_{\text{air}}$		3	3·5	5·7	23
$0{\cdot}2(w/N)_{\text{oxygen}} + 0{\cdot}8(w/N)_{\text{nitrogen}}$		3·1	3·5	5·3	17

Losses of energy in mixtures. If w_a and w_b are the mean rates at which electrons, moving in a steady state of motion with the same mean energy $\bar{Q}$, lose energy in each of two gases a and b separately it is usually assumed (method of mixtures) that, in a mixture comprising x parts of gas a and y parts of gas b ($x + y = 1$) in a state of motion with the same energy $\bar{Q}$, electrons lose energy at the rate:

$$w_m = xw_a + yw_b$$

Since the chief contributions to a rate of energy loss w come from those electrons of the group that collide with energies Q appreciably greater than $\bar{Q}$ it is to be anticipated that, when $\bar{Q}$ is given, w will also depend upon the manner in which the energies Q are distributed about $\bar{Q}$.

The assumption that $w_m = xw_a + yw_b$ implies the further assumption that when $\bar{Q}$ is specified the energies Q are distributed in precisely the same manner about $\bar{Q}$ whether the electrons move in the mixture or in the gases a and b separately. There is no reason to suppose that such an assumption is valid in general.

The behaviours of electrons in nitrogen, oxygen and air (Table 4) serve to illustrate the failure of the method of mixture in some circumstances and its validity in others.

It is evident that the formula for mixtures does not apply when q exceeds 6 but does so when q is less than 6. It may be inferred that when $q < 6$ then the manner in which the energies Q are distributed about $\bar{Q}$ is essentially the same in air, oxygen and nitrogen and, presumably, according to Maxwell's law.

Unfortunately no information about electronic motion in oxygen when $q < 4{\cdot}7$ is available, but when $q = 4$, $(w/N)_{\text{air}}$ and $(w/N)_{\text{nitrogen}}$ have become almost equal. This behaviour is consistent with the suggestion made above that $(w_r/N)_{\text{nitrogen}}$ and

$(w_r/N)_{\text{oxygen}}$ are roughly equal when $q < 4$. It is assumed therefore in what follows that formulae (19) is valid not only for the motions of electrons in nitrogen but in air also. According to equations (19) and (21), when $q \to 1$

$$\left.\begin{aligned} & w/N \to B(Q - Q_0) \\ \text{with} \quad & \\ & B = \alpha\beta Q_0^{-3/2} \exp(-\beta/Q_0) + 3{\cdot}78Q_0 \\ \text{and} \quad & \\ & \alpha = 7{\cdot}29 \times 10^{-18}; \qquad \beta = 4{\cdot}41 \times 10^{-14} \end{aligned}\right\} \tag{26}$$

Thus, B is temperature dependent through Q_0; for instance

$T°\text{K}$	288	196
$B \times 10^{11}$	1·07	1·34

The interaction of radio waves

The theory of the interaction of radio waves is based upon an equation for the fluctuation of the mean energy of agitation Q of the electrons about a mean value $\bar{Q}$ under the influence of the modulated disturbing wave. This basic equation is

$$\frac{dQ}{dt} + R = w(t)$$

where $w(t)$ is the power supplied in the mean to an electron by the disturbing wave and R is the mean rate at which an electron with mean energy Q loses energy in collisions with the molecules of the atmosphere.

This equation is usually solved in terms of the assumption that $R = C(Q - Q_0)$, where C is a constant. From what has preceded it is evident that this assumption is legitimate when $Q \to Q_0$ and that $C = BN$.

Measurements of radio-wave interaction lead to values of C which were usually found to be within the limits $10^3 < C < 3 \times 10^3$, with values in the vicinity of 2×10^3 of common occurrence (HUXLEY, 1952, where C is written as $G\nu$).

The measured heights of the seat of the cross-modulation usually lay between the values 82 and 90 km above the ground. According to the A.R.D.C. model atmosphere (MINZNER and RIPLEY, 1956) the atmosphere within this range of heights has a composition essentially the same as that at the ground, a temperature of $T = 196°\text{K}$ and a scale height of about 5·92 km.

Thus, the appropriate value of B is $1{\cdot}34 \times 10^{-11}$ and the value of the molecular concentration associated with a measured value of C is $N = C/B$.

The following table gives the values of N associated with the values 10^3, 2×10^3 and 3×10^3 of C that cover the range of measured values of C. The heights in the A.R.D.C. model atmosphere at which these values of N are found are also shown.

VOL. 16 1959

These heights evidently fall within the range of measured heights from 82 to 90 km.

$T = 196°\text{K};\ B = 1{\cdot}34 \times 10^{-11}$

C	10^3	2×10^3	3×10^3
$N = C/B$	$7{\cdot}45 \times 10^{13}$	$1{\cdot}49 \times 10^{14}$	$2{\cdot}23 \times 10^{14}$
A.R.D.C. height (km)	90·7	86·6	84·2

It may be concluded that the A.R.D.C. model atmosphere is consistent with the measured properties of radio-wave interaction. A more systematic comparison could be effected if accurate measurements of C and the corresponding heights h could be made for two or more heights. Let C_1, N_1 and C_2, N_2 correspond to h_1 and h_2. Then $C_1/C_2 = N_1/N_2$ so that the scale height H is deducible. From H the temperature T is known and from T the value of B and therefore those of N_1 and N_2 also.

The conductivity matrix of the lower ionosphere

The following expressions have been given (HUXLEY, 1957b) for the a.c. conductivity of a weakly ionized gas containing free electrons in the presence of a magnetic field **B**.

The electric vector has components $E_x = X_0 \cos(pt + \alpha)$, $E_y = Y_0 \cos(pt + \beta)$ and $E_z = Z_0 \cos(pt + \gamma)$, and the complex current density $\mathbf{J}(J_x, J_y, J_z)$ is given by

$$\mathbf{J} \equiv \begin{pmatrix} J_x \\ J_y \\ J_z \end{pmatrix} = \|\sigma^+\| \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \exp(ipt) + \|\sigma^-\| \begin{pmatrix} X^* \\ Y^* \\ Z^* \end{pmatrix} \exp(-ipt)$$

where $X = X_0 \exp(i\alpha)$, $Y = Y_0 \exp(i\beta)$ and $Z = Z_0 \exp(i\gamma)$,

and $$\|\sigma^+\| = \begin{Vmatrix} \sigma_{xx}^+ & \sigma_{xy}^+ & 0 \\ \sigma_{yx}^+ & \sigma_{yy}^+ & 0 \\ 0 & 0 & \sigma_{zz}^+ \end{Vmatrix}, \|\sigma^-\| = \begin{Vmatrix} \sigma_{xx}^- & \sigma_{xy}^- & 0 \\ \sigma_{yx}^- & \sigma_{yy}^- & 0 \\ 0 & 0 & \sigma_{zz}^- \end{Vmatrix}$$

where,

$$\left.\begin{aligned} \sigma_{xx}^+ = \sigma_{yy}^+ = i\sigma_{xy}^+ = -i\sigma_{yx}^+ = (ne^2/6m) \cdot \overline{c^{-2}\frac{d}{dc} \cdot \frac{c^3}{\nu - i(\omega - p)}} \\ \sigma_{xx}^- = \sigma_{yy}^- = i\sigma_{xy}^- = i\sigma_{yx}^- = (ne^2/6m) \cdot \overline{c^{-2}\frac{d}{dc} \cdot \frac{c^3}{\nu - i(\omega + p)}} \\ \sigma_{zz}^+ = (ne^2/6m)\overline{c^{-2}\frac{d}{dc} \cdot \frac{c^3}{\nu + ip}} \end{aligned}\right\} \quad (27)$$

and $\omega = -Be/m$ with **B** directed along $+Oz$ so that ω is positive for electrons for which e is a negative quantity; n is the number of electrons in unit volume.

The corresponding matrix of the conductivity implicit in the usual magneto-ionic theory of radio-wave propagation in the ionosphere is:

$$\left.\begin{aligned}
&\sigma_{xx}^{+} = \sigma_{yy}^{+} = i\sigma_{xy}^{+} = -i\sigma_{yx}^{+} = (ne^2/2m)\,\frac{1}{\bar{\nu} - i(\omega - p)} \\
&\sigma_{xx}^{-} = \sigma_{yy}^{-} = i\sigma_{xy}^{-} = -i\sigma_{yx}^{-} = (ne^2/2m)\,.\,\frac{1}{\bar{\nu} - i(\omega + p)} \\
&\sigma_{zz}^{+} = (ne^2/2m)\,.\,\frac{1}{\bar{\nu} + ip}
\end{aligned}\right\} \qquad (28)$$

These expressions (28) are the same as the more rigorous expressions (27) in two special instances only:

(a) when $\nu = 0$.

(b) when $\nu = NAc$ is a constant independent of c

This implies that $A \propto 1/c$, that is to say, that the electrons are repelled by the molecules by a force varying inversely as the fifth power of the distance.

The studies of the motions of electrons in nitrogen discussed in this paper show however that case (b) is not applicable to the lower ionosphere where $A \propto c$ and not to c^{-1}. Errors, by factors of several units, may arise when the normal magneto-ionic expressions (28) are used to interpret measurements of radio propagation in the lower ionosphere where the effects of collisions are important.

Acknowledgements—The author is indebted to Drs. R. W. CROMPTON and BARBARA HALL for making their experimental measurements available to him, and to Dr. CROMPTON for helpful discussions.

REFERENCES

CROMPTON R. W., HUXLEY L. G. H. and SUTTON D. J.	1953	*Proc. Roy. Soc.* A **218,** 507.
HUXLEY L. G. H.	1952	*Nuovo Cim.* **9,** (Series 9, Supplement)
HUXLEY L. G. H.	1956	*Aust. J. Phys.* **9,** 1, 44.
HUXLEY L. G. H.	1957a	*Aust. J. Phys.* **10,** 1, 118.
HUXLEY L. G. H.	1957b	*Aust. J. Phys.* **10,** 2, 240.
MINZNER R. A. and RIPLEY W. S.	1956	*The A.R.D.C. Model Atmosphere.* Air Force Surveys in Geophysics, No. 86.

VOL. 16 1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 59 to 66. Pergamon Press Ltd. Printed in Northern Ireland

Type-*B* aurora in the Antarctic*

J. M. MALVILLE
Yerkes Observatory, University of Chicago

(*Received* 20 *October* 1958)

Abstract—Spectrograms obtained in the Antarctic during 1957 of aurorae with red lower borders are described. The outstanding feature of this type of aurora is an enhancement of the first positive system of N_2. A reaction of mutual neutralization of the N_2^+ molecule by the O_2^- negative ion is proposed to account for the observed enhancement. Assuming optimum values for rate coefficients and an altitude of 70 km, mutual neutralization will proceed with a time constant of the order of 1 sec, consistent with the observed short life time of type-*B* aurora. The observed enhancements of the first negative system of O_2^+ and of sodium "*D*" are also postulated as resulting from the penetration of an energetic beam of electrons to the 70–90 km level of the *D*-region. A monoenergetic beam of electrons and protons of 240 keV is compatible with the elevations of homogeneous arcs and type-*B* aurora. Other possible excitation mechanisms for type *B* are discussed.

1. INTRODUCTION

Two types of red coloured aurorae are distinguished: type-*A* possessing a diffuse red colour and type-*B* having a sharp red lower border. As both these types are short lived, little adequate information concerning their spectral nature has been available. According to VEGARD (1938) type-*A* aurora is the result of enhancement of the 6300/6364 (O*I*) doublet and occurs most frequently at time of sunspot maximum. Type-*B* on the other hand appears to be the result of a low altitude enhancement of the first positive band system of the nitrogen molecule and according to VEGARD is most frequent at sun spot minimum. Recently HUNTEN (1955) using a scanning spectrophotometer has found that the first negative system of the ionized oxygen molecule is similarly enhanced in type-*B*. HUNTEN also finds evidence for the enhancement of the sodium *D*-line and a suppression of the red oxygen doublet in type-*B*.

While a project scientist for the Auroral and Airglow Panel of the U.S. National Committee of the IGY the author made auroral observations during the Antarctic winter of 1957 at Ellsworth Station in the Weddell Sea (geomagnetic latitude, 67° S). Photographs with the IGY all-sky camera and low dispersion spectra were obtained of type-*B* aurora. This paper discusses the results of these observations and suggests an excitation mechanism to explain the details of the type-*B* spectrum.

2. OBSERVATIONS

Type-*B* aurora were observed during sixteen nights of the 1957 Antarctic winter. Generally concurrent with the break-up of a quiet arc into active rayed aurora, type-*B* occurred predominantly before midnight and possessed rapid east–west motions with directions usually dependent upon the local time (west before magnetic midnight, east after magnetic midnight). Measurements of the all-sky camera photographs give east–west velocities of the order of 300 m/sec, in

* The research reported in this paper was supported in part by the Geophysics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command, under Contract AF 19(604)–3044.

agreement with previous measurements of velocities of active aurorae (MEINEL and SCHULTE, 1953).

Fig. 1 shows spectrograms of type-*B* obtained with the IGY patrol spectrograph at Ellsworth Station. The slit of the spectrograph, 2° wide and 180° long, was aligned parallel to the geomagnetic meridian. Luminosity curves of the principal radiations are given in Fig. 2 for the spectrogram of 15–16 August. A quiet homogeneous arc remained in the southern sky between 0018 UT and 0050

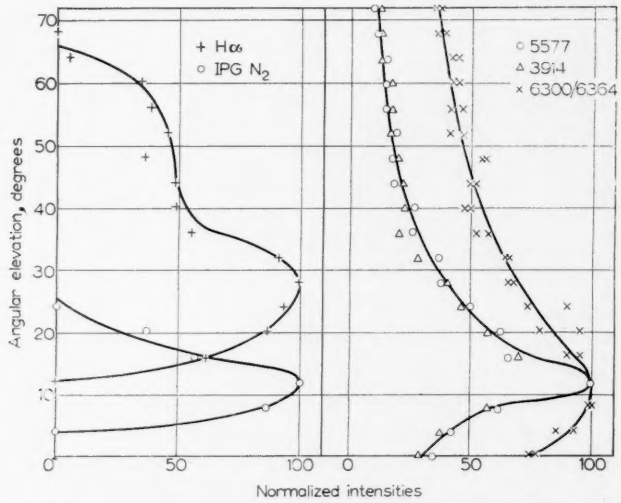


Fig. 2. Luminosity curves of the spectrogram of 15–16 August. All intensities have been normalized to 100 at the elevation of maximum intensity.

UT at an elevation of 20°; at 0055 UT a minor break-up occurred with the appearance of a rayed band with an international brightness coefficient of II (SEATON, 1954); and at 0105 UT the true break-up occurred with the appearance of a brightness III type-*B* aurora. The following prominent features of the spectrograms are noted:

(a) The maximum of H-α appears higher in elevation than the other auroral emissions.

(b) The first positive system of N_2 is enhanced with a maximum width of 5°, with no similar enhancement of the second positive system.

(c) The second positive system decreases in intensity with altitude less rapidly than does the first positive system.

(d) The intensity ratio ($\lambda 5577/\lambda 3914$) is constant with elevation with both the radiations peaked at the elevation of the N_2 first positive group maximum.

(e) The red oxygen doublet shows the least variation with elevation, but is nevertheless peaked at the elevation of the maximum of the first positive group.

3. INTERPRETATION OF THE SPECTRA

The separation in elevation between H-α and the first positive system may be explained as the result of an altitude difference between the original homogeneous

VOL. 16 1959

TOI. 16 1959

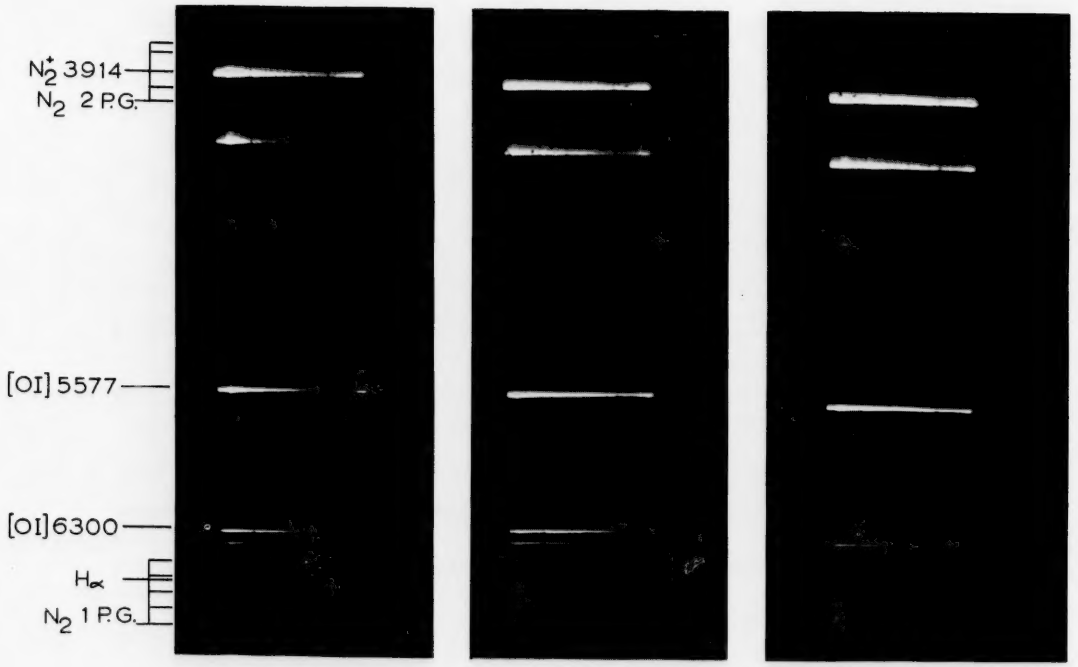


Fig. 1. Spectrograms of type-*B* aurora obtained at Ellsworth Station during 1957. The top of each spectrogram corresponds to the northern horizon, the bottom to the southern horizon. (a) 0208–0228 UT, 5–6 June; (b) 0115–0235 UT, 19–20 June; (c) 0018–0114 UT, 15–16 August.

VOL.
16
1959

arc and the type-B aurora. Type-B aurorae appear much lower in the atmosphere than the other auroral forms. HARANG (1951) reports parallactic photographic measurements of type-B which show a systematic decrease in altitude of 20 km at the time of appearance of a red lower border. We can thus picture the H-α radiation as arising from the quiet homogeneous arc which existed before the break-up; at the time of the 0105 UT break-up the altitude of the aurora dropped to the 70–90 km level and presumably as the result of some process unique to this layer of the atmosphere the first positive system of N_2 was enhanced. As the H-α radiation appears to differ from the type-B not only in time but in altitude it seems likely that some agency other than primary auroral protons must be responsible for the excitation of this form of aurora.

Probably at least two separate excitation processes are active on the N_2 molecule: one which excites only the first positive system in the 70–90 km layer and the more common one (probably electron impact) which excites both the first and second positive system.

4. Mechanism for the Excitation of the N_2 First Positive System

Any mechanism which is proposed to account for the unique feature of type-B aurora must meet the following requirements:

(a) Provide a strong enhancement of the first positive N_2 system.
(b) Have a decay time of 1 sec.
(c) Must be unique to the level of the atmosphere below 90 km.
(d) Must be compatible with high auroral luminosity and consequently with a high degree of ionization.

The most obvious choice for an excitation mechanism operative only in the D-region is one involving negative ions. At night in the D-region high concentrations of the negative ions O^- and O_2^- are formed by the reactions:

$$O + e \rightarrow O^- + h\nu \tag{1}$$

$$O_2 + e \rightarrow O_2^- + h\nu \tag{2}$$

MITRA (1946) and GHOSH (1946) have proposed that in aurora the N_2 first positive sytem is excited by a series of reactions starting with the ionization of N_2 leaving the N_2^+ molecule in the upper state of the first negative system, $B^2\Sigma_u^+$. After radiating, the N_2^+ molecule in its ground state, $X^2\Sigma_g$, undergoes mutual neutralization with the negative ion O^- leaving the N_2 molecule in the upper state of the first positive system:

$$N_2^+ + O^- \rightarrow N_2(B\,^2\Pi) + O \tag{3}$$

As this mechanism is energetically incapable of exciting the upper level of the second positive system it is, although inadequate in the case of normal aurora, quite satisfactory for type-B. At the 70–90 km level, however, because the ratio $O_2/O \sim 10^3$ (BATES and WITHERSPOON, 1952) a process involving the negative oxygen molecular ion would be more significant and would proceed more rapidly. Consequently a process involving the mutual neutralization of N_2^+ in its ground

state by O_2^- is capable of accounting for the observed enhancement of the first positive N_2 system relative to the second positive system:

$$N_2^+(X\ ^2\Sigma_g^+) + O_2^-(^2\Pi_g) \rightarrow N_2(B\ ^3\Pi_g) + O_2(^3\Sigma_u) \qquad (4)$$

The energies and electron configurations of the appropriate states are given in Table 1.

Table 1

Ion	State	Energy (eV)	Electron configuration
N_2^+	$X\ ^2\Sigma_g^+$	15·6	$(\sigma_g 2s)^2(\sigma_u 2s)^2(\pi_u 2p)^4(\sigma_g 2p)$
N_2	$C\ ^3\Pi_u$	12·14	$(\sigma_g 2s)^2(\sigma_u 2s)(\pi_u 2p)^4(\sigma_g 2p)^2(\pi_g 2p)$
N_2	$B\ ^3\Pi_g$	9·8	$(\sigma_g 2s)^2(\sigma_u 2s)^2(\pi_u 2p)^4(\sigma_g 2p)(\pi_g 2p)$
O_2	$^3\Sigma_u$	5	$(\sigma_g 2s)^2(\sigma_u 2s)^2(\pi_u 2p)^3(\sigma_g 2p)^2(\pi_g 2p)^3$
O_2^-	$^2\Pi_g$	1	$(\sigma_g 2s)^2(\sigma_u 2s)^2(\pi_u 2p)^4(\sigma_g 2p)^2(\pi_g 2p)^3$

As with the O^- reaction (4) is resonant for the upper level of the first positive system. A consideration of the electron configurations likewise gives a low probability for the excitation of the $C^3\Pi_u$ state from the ground state of N_2^+. An internal rearrangement of electrons is required in addition to electron exchange in passing from the $X^2\Sigma_g^+$ state to the $C^3\Pi_u$ state. On the other hand only the exchange of a $\pi_g 2p$ electron is required to produce the $B^3\Pi_g$ state from the ground state of N_2^+.

Reaction (4) will be confined to the region of the atmosphere where the ratio of negative ions to electrons, λ, is equal to or greater than unity. At night with the absence of photo-ionization we can approximate the equilibrium value of λ as

$$\lambda \sim \frac{\beta n(O_2)}{\kappa n(O)} \qquad (5)$$

if we assume that

$$O_2^- + O \rightarrow O_3 + e \qquad (6)$$

is the most significant electron detachment process, where β is the coefficient of electron attachment and κ the coefficient of electron detachment. At the 70 km level $\{n(O_2)/n(O)\} = 1{\cdot}1 \times 10^4$ and at 100 km $\{n(O_2)/n(O)\} = 10^{-1}$ (BATES and WITHERSPOON, 1952). Taking the ratio of β/κ to vary between 10 and 10^{-1} gives the following likely ranges of λ for night-time equilibrium:

$$1{\cdot}1 \times 10^3 \leq \lambda_{70} \leq 1{\cdot}1 \times 10^5$$

$$1{\cdot}1 \times 10^{-2} \leq \lambda_{100} \leq 1{\cdot}1$$

The observational results of MITRA (1957) indicate a value of $\lambda_{80} = 40$. Somewhere between 90 and 110 km λ becomes much less than unity, mutual neutralization becomes insignificant and dissociative recombination:

$$N_2^+ + e \rightarrow N^* + N^* \qquad (7)$$

becomes the important reaction involving the N_2^+ ion.

VOL. 16 1959

Thus in the series of reactions consisting of the ionization of N_2, the subsequent production of ions in the ground state of N_2^+, and the mutual neutralization of N_2^+ by O_2^- we appear to have a reasonable mechanism for explaining the emission features of type-B aurora. But before asserting this as a likely mechanism we must demonstrate that it will proceed with a time constant of the order of 1 sec as required by the short life time of type-B.

If we take N_e, N_-, N_+, and N_r to be respectively the concentrations of electrons, negative ions, positive ions, and neutral particles, and β, κ, α_i, and α_d to be respectively the coefficients of electron attachment, electron detachment, mutual neutralization and dissociative recombination, then we may write the following three coupled equations:

$$\frac{dN_e}{dt} = -\beta N_e N_r + \kappa N_- N_r - \alpha_d N_e N_+ \tag{8}$$

$$\frac{dN_-}{dt} = \beta N_e N_r - \kappa N_- N_r - \alpha_i N_- N_+ \tag{9}$$

$$\frac{dN_+}{dt} = -\alpha_i N_- N_+ - \alpha_d N_e N_+ \tag{10}$$

Fig. 3 shows the result of the numerical integration of these three equations. Using optimum values for the rate coefficients, the concentration of negative ions does indeed decrease to $1/e$ of its maximum value in the order of 1 sec (approximately 5 sec in Fig. 3a and 1·3 sec in Fig. 3c). However with the decrease of either the rate coefficients or of the density of neutral particles the decay time of N_- rapidly becomes too long to be compatible with the observed phenomenon.

The enhancement of the negative system of O_2^+ in type-B as observed by Hunten follows as a natural consequence of the high concentration of O_2 in the D-region. According to Bates and Witherspoon (1952) the ratio of the concentration of O_2 at 70 km to that at 110 km equals 8×10^4. However the situation is not as simple with the observed enhancement of the sodium D-line. Although type-B arise in the atmospheric region of the maximum sodium concentration, because of the low density of free sodium atoms, excitation by straightforward electron impact can not be responsible for the Na D-enhancement. It is possible that a sodium compound such as NaO or NaO_2 is ionized by incident electrons and is in turn involved in dissociative recombination leaving the sodium atoms in an excited state, thus:

$$NaX^+ + e \rightarrow Na^* + X \tag{11}$$

Electron densities of the order of 10^7 per cm^3 are not unreasonable from estimates derived from auroral intensities (Seaton, 1954) and from auroral radio reflections (Currie *et al.*, 1953). Assuming that one ionization of N_2 out of fifty leads to a λ3914 quantum, we are able to set an approximate lower limit on type-B associated electron densities. The rate of electron production and removal is governed by the equation:

$$\frac{dN_e}{dt} = q - \alpha_{rec} N_e^2 \; ; \; q = \frac{50}{L} \times f_{3914} \times Q(3914) \tag{12}$$

where $f_{3914} = 10$ = Seaton's ultra-violet correction factor (SEATON, 1954), $L = 1$ km = approximate effective auroral path length, and $Q(3914)$ = number of 3914 quanta emitted per sec per cm² column. When $dN_e/dt = 0$ then $N_e^2 > q/\alpha_{rec}$. The effective recombination coefficient increases rapidly with decreasing altitude, but it is doubtful that at 70 km α_{rec} is greater than 10^{-6} cm³/sec. Thus with an aurora with an international brightness coefficient of III and a $Q(3914)$ of 10^{11} quanta/sec per cm², $N_e > 2 \times 10^7$ cm^{-3}.

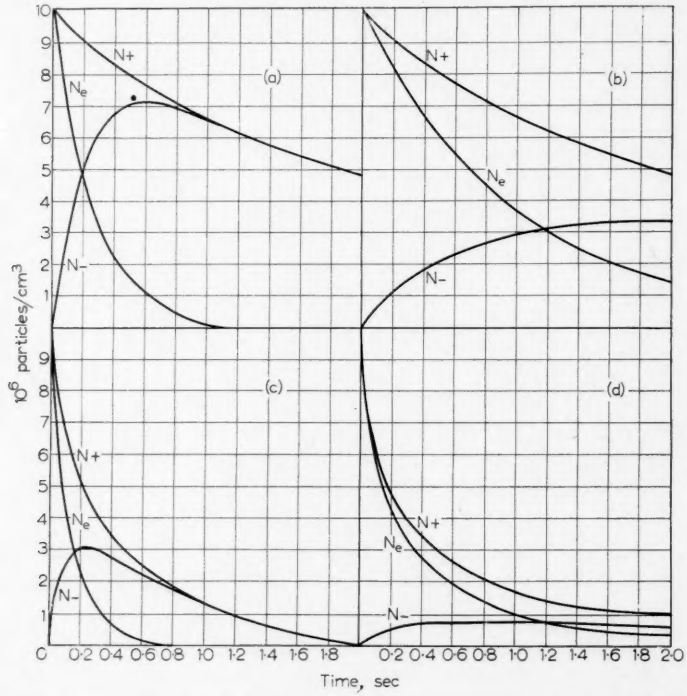


Fig. 3. Curves derived from the numerical integration of equations (8), (9) and (10) for an altitude of 70 km and the initial conditions: $t = 0$, $N_- = 0$, $N_e = N_+ = 10^7$/cm³. The following values for the rate coefficients were used: (a) $\beta = 10^{-14}$ cm³/sec, $\alpha_d = \alpha_i = 5 \times 10^{-8}$ cm³/sec; (b) $\beta = 1{\cdot}5 \times 10^{-15}$ cm³/sec, $\alpha_d = \alpha_i = 5 \times 10^{-8}$ cm³/sec; (c) $\beta = 10^{-14}$ cm³/sec, $\alpha_d = \alpha_i = 5 \times 10^{-7}$ cm³/sec; (d) $\beta = 1{\cdot}5 \times 10^{-15}$ cm³/sec, $\alpha_d = \alpha_i = 5 \times 10^{-7}$ cm³/sec.

Note: if β is held constant then passing from curves (a) to (b) or from (c) to (d) shows the effect of increasing altitude

5. PRODUCTION OF IONIZATION

The problem remains of explaining the apparent sudden burst of ionization at the 70–90 km level of the atmosphere. As it appears that primary protons are not directly involved with the type-*B* phenomenon, the problem is related to that of the sudden break-up of the aurora, the transition from the quiet stage to the active one of predominantly electronic excitation. Evidence for the existence of electrons of energy up to 100 keV near the 80 km level has recently been obtained by VAN ALLEN and associates (MEREDITH *et al.*, 1955) in the auroral zone, and by WINCKLER (1958) and co-workers at Minneapolis on balloon flights into visible aurora.

VOL. 16 1959

For electrons to penetrate to the 70 km level in the atmosphere requires energies of 240 keV (BATES, 1954; BATES and GRIFFING, 1953). Protons with energies of 240 keV would penetrate to 110 km, the level of the atmosphere where homogeneous arcs are most frequently observed. Thus a mono-energetic beam of electrons and protons would produce an effect compatible with the observed phenomenon. As protons of 240 keV have velocities a third of that of electrons of the same energy, electrons cannot have been members of the original beam producing the quiet arc preceding the type-*B*. It is possible, however, that a localized field is set up in the earth's atmosphere such as is proposed by several recent auroral theories (CHAMBERLAIN, 1957).

Mass motions in the *D*-region might also be responsible as a supplementary mechanism for exciting the first positive system of N_2. BERNARD (1954) has proposed that the enhancement of the first positive N_2 system is the result of motions of the atmosphere with respect to a moving electrical discharge. BERNARD and PEYRON (1948) have observed such an enhancement in a laboratory discharge tube when there was a rapid circulation of nitrogen.

BYRAN *et al.* (1957) have also observed in the laboratory an enhancement of the first positive group in a nitrogen afterglow following the discharge of a microwave cavity. They observed after a strong discharge a red afterglow differing significantly from the longer lived Lewis–Rayleigh nitrogen afterglow. The first positive group was enhanced in the red afterglow, but not in the "blue" afterglow following a weaker discharge. It is possible that a similar mechanism is responsible for the observed enhancement in type-*B* aurora as a strong, localized, horizontal electric field could conceivably be created by the motion of ionized particles of the *D*-region normal to the magnetic lines of force.

However, in both these mechanisms one must rely upon the elusive agency of *D*-region winds and associate these winds with the time variation of the auroral phenomenona. It is more reasonable to picture corpuscular radiation as the primary excitation mechanism for the type-*B* phenomenon, and to relegate atmospheric motions to the role of an effect rather than a cause.

Acknowledgement—The author gratefully acknowledges Dr. J. W. CHAMBERLAIN for his helpful interest and assistance in this problem.

REFERENCES

BATES D. R.	1954	*The Earth as a Planet*, (Edited by G. P. Kuiper) University of Chicago Press.
BATES D. R. and GRIFFING G.	1953	*J. Atmosph. Terr. Phys.* **3**, 212.
BATES D. R. and WITHERSPOON A. E.	1952	*Mon. Not. R. Astr. Soc.* **112**, 101.
BERNARD R. G.	1954	*Proc. Conf. Auroral Phys.* Air Force Cambridge Research Center, Geophysical Research Papers, No. 30.
BERNARD R. G. and PEYRON M.	1948	*C.R. Acad. Sci. Paris* **227**, 1151.
BYRAN R. B., HOLT R. B. and OLDENBERG O.	1957	*Phys. Rev.* **106**, 83.
CHAMBERLAIN J. W.	1957	*Astrophys. J.* **126**, 245.
CURRIE B. W., FORSYTH P. A. and VAWTER F. E.	1953	*J. Geophys. Res.* **58**, 179.

GHOSH S. N.	1946	*Proc. Nat. Inst. Sci. India* **17,** 405.
HARANG L.	1951	*The Aurorae.* John Wiley, New York.
HUNTEN D. M.	1955	*J. Atmosph. Terr. Phys.* **7,** 141.
MEINEL A. B. and SCHULTE D. H.	1953	*Astrophys. J.* **117,** 454.
MEREDITH L. H., GOTTLEIB M. B. and VAN ALLUN J. A.	1955	*Phys. Rev.* **97,** 201.
MITRA A. P.	1957	*J. Atmosph. Terr. Phys.* **10,** 153.
MITRA S. K.	1946	*Nature, Lond.* **157,** 692.
OMHOLT A.	1955	*J. Atmosph. Terr. Phys.* **7,** 345.
SEATON M. J.	1954	*J. Atmosph. Terr. Phys.* **4,** 285.
VEGARD L.	1938	*Nature, Lond.* **141,** 200.
WINCKLER J. R., PETERSON L., ARNOLDY R. and HOFFMAN R.	1958	*Phys. Rev.* **110,** 1221.

VOL.
16
1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 67 to 83. Pergamon Press Ltd. Printed in Northern Ireland

New methods and some results concerning true ionospheric height calculations*

W. BECKER
Max-Planck-Institut für Aeronomie, Institut für Ionosphären-Physik
in Lindau am Harz, Germany

(*Received* 17 *October* 1958)

Abstract—The methods can be listed as follows:
(a) Optical-graphical comparison method, comparing, for certain layer-types, calculated ordinary and extraordinary $h'(f)$-traces with observed $h'(f)$-records. It is possible to determine the best fitting layer-type and its parameters within 8 min.
(b) General method, applicable to any monotonic $h'(f)$-trace. This method is based on RYDBECK's solution of the respective integral equation. A 10 point $h'(f)$-reduction according to this method takes 2 hr.
(c) Correction method, combining the methods (a) and (b). Thus the time necessary for a 10 point $h'(f)$-reduction can be reduced to 5/4 hr.
The following results are reported:
(a) Good agreement between actual layer shape and parabolic $N(h)$-distribution for the undisturbed night-time F-layer on 8 August 1957.
(b) Periodic movements ($T \sim 1$ hr, amplitude $\sim$8 km) of the F-layer during a magnetically undisturbed night, 7–8 August 1957.
(c) Sudden ascent of the F-layer during a "Polar Sudden Commencement" ($\Delta h \sim 60$ km within $\frac{1}{2}$ hr on 15–16 September 1956.
(d) Existence of an $N(h)$-minimum above the normal E-layer above Lindau, Germany, on 20 August 1957, 1545 hours.

INTRODUCTION

MORE than sixty papers should be referred to before discussing this problem as was shown by BECKER (1955) and more recently by KING (1957). It is not the intention of this paper to supplement these summaries but to report on own research work in this field. A subdivision of this report into three parts is useful:

Section (1) describes the comparison method now in use at this station (Lindau/Harz, Germany)†.

Section (2) deals with our general methods, applicable to any monotonic $h'(f)$-trace.

Section (3) gives some results.

1. COMPARISON METHOD

This section describes a comparison method for routine determinations of vertical electron density distributions in ionospheric layers from their sweep-frequency records.

1.1. *Determinations from ordinary* h'(f)*-traces*

It is well known that the earth's magnetic field considerably influences the

* A short outline of this paper appeared in "Research Report EE 361" of Cornell University, School of Electrical Engineering, Ithaca, New York, 30 January 1958.

† A preliminary description of this comparison method was given by the author in 1956. A full elaboration of this method is presented in this hereport. T respective numerical data have been published [BECKER W. (1959) *Arch. Elektr. Übertr.* **13,** 49].

time of travel of pulses of radio energy within an ionospheric layer. Fig. 1 demonstrates this effect. The symbols used in Fig. 1 agree with those of SHINN (1954) and BECKER (1957). G measures the ratio of the ordinary vertical group velocities for 0 and 67·1°, angles of dip (J) of the earth's magnetic field. The plasma frequency is denoted by f_0 . $t = \sqrt{(1 - f_0^2/f^2)}$. Y is the ratio f_H/f, where f_H is the gyrofrequency and f the operating frequency.

The influence of electron collisions on the vertical group velocities is neglected. It is this influence of the earth's magnetic field that allows only numerical solutions

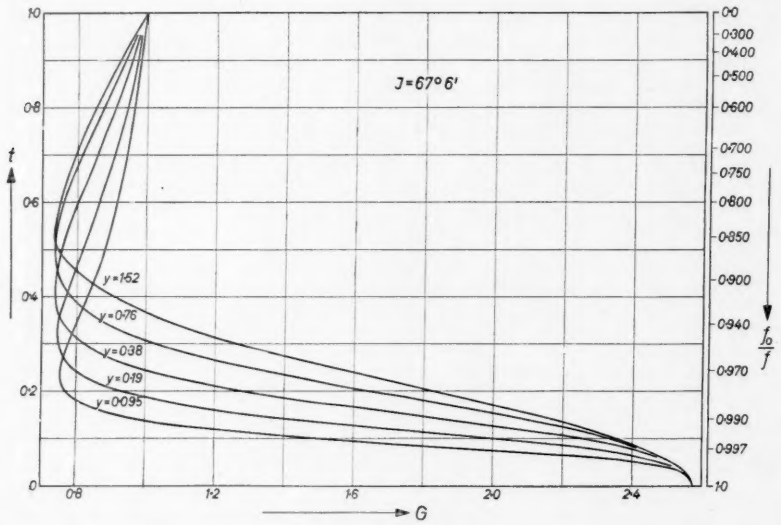


Fig. 1. G measures the ratio of the ordinary vertical group velocities for 0 and 67·1° angles of dip, J, of the earth's magnetic field. $t = \sqrt{(1 - f_0^2/f^2)}$; f_0 = plasma frequency, f = operating frequency; $y = f_H/f$; f_H = gyrofrequency.

of the respective integral equation. But even these calculations are so complicated that an electronic computing machine, when available, does not allow a comprehensive reduction technique (KELSO, 1954; BUDDEN, 1954). Therefore, another method was sought. A method was derived based upon a graphical comparison of calculated and observed ordinary virtual heights of an ionosphere with one or more layers. In order to be independent of a special layer type within certain limits, of course, three types are utilized: (a) an Epstein, (b) a cosine and (c) a parabolic layer (Fig. 2). By equivalent we mean that these layers have equal maximum electron densities and equal electron content. Thus, it is possible to express the electron content of each layer and all heights in units of half the layer thickness Y_m of the equivalent parabolic layer. Now, it is not the intention of this proposal to compare the whole sweep-frequency patterns with each other; that would be too circumstantial, since these patterns are not only functions of Y_m but also of the ratio f_H/f_c, as Fig. 3 shows. Each curve belongs to one of the assumed layer types; f_H/f_c is taken equal to 0·3724. Only a series of specially selected, equally spaced virtual heights (h'_a–h'_0) belonging to certain values of

VOL. 16 1959

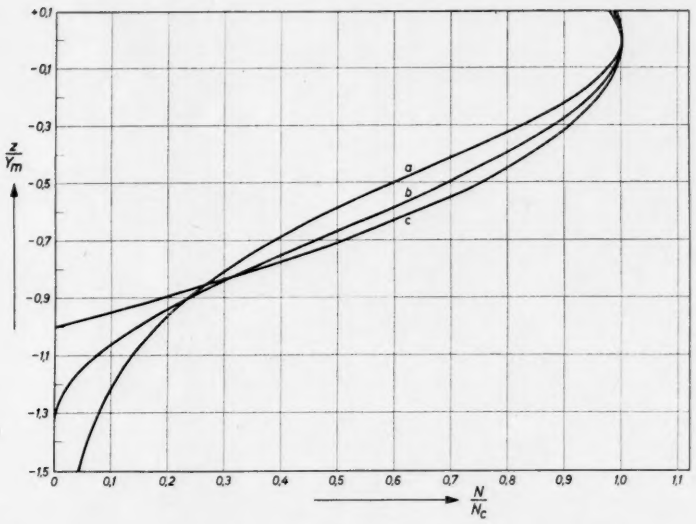


Fig. 2. Electron density profiles representing an (a) Epstein, (b) a cosine and (c) a parabolic layer shape; equal maximum densities and equal total contents are assumed. Y_m measures the half layer thickness of the parabolic profile.

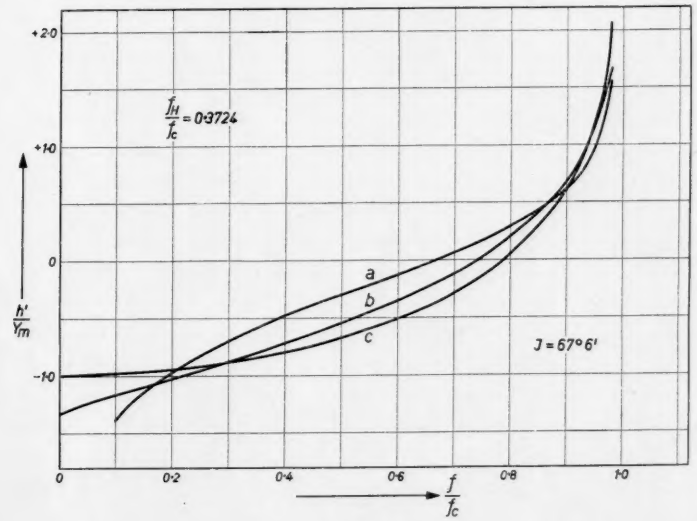


Fig. 3. Calculated ordinary $h'(f)$-curves for an (a) Epstein, (b) a cosine and (c) a parabolic layer shape. J = angle of dip of the earth's magnetic field; f_c = ordinary critical frequency; f_H = gyrofrequency.

$f/f_c = \varkappa$ are compared. Fig. 4 represents these $\varkappa$-values and the respective virtual height parameters for a cosine layer. The curves show how the reading frequency must be altered in order to always obtain constant reflection heights for different critical frequencies.

Curve f represents the dependence on critical frequency of the Booker–Seaton's

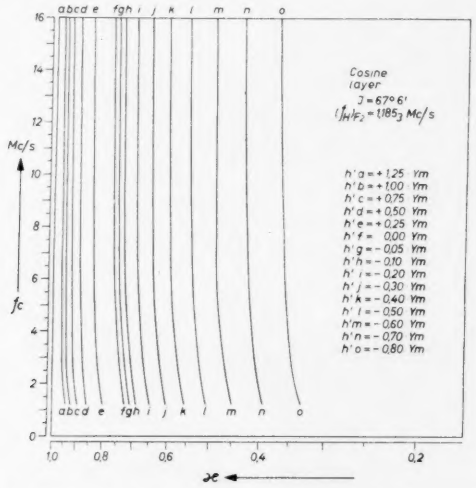
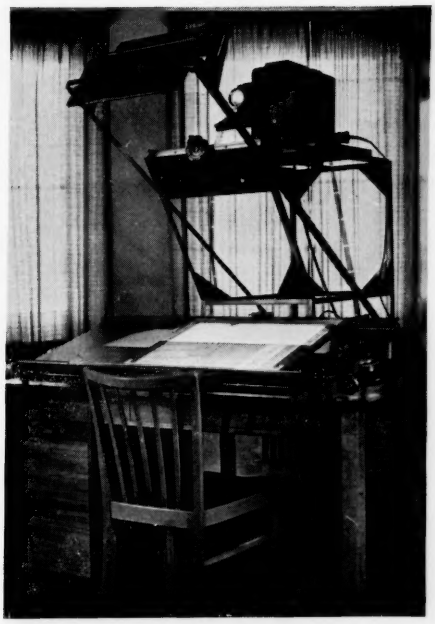


Fig. 4. Ordinary critical frequency f_c versus frequency of observation $f(\varkappa = f/f_c)$ for constant ordinary virtual heights $h'_a - h'_0$ of a cosine electron density profile. Y_m measures the equivalent parabolic half layer thickness, J the angle of dip of the earth's magnetic field and f_H the gyrofrequency.

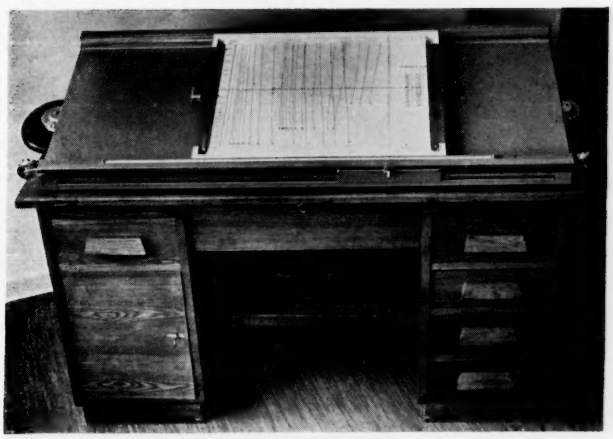
factor 0·7925 for a cosine layer (it is 0·834 for a parabolic layer). Now, the comparison can be realized in the following way: one of the three sheets of paper with the respective $\varkappa$-curves for the different layer types is attached to the slider of Fig. 5. The ionospheric sweep-frequency record is optically projected on the frame in such a way that the logarithmic scale of the record and the logarithmic $\varkappa$-scale correspond to each other. Then, the vertical pointer of the frame is moved to the ordinary critical frequency and the horizontal pointer of the slider is moved to the same value at the slider's f_c-scale. Finally, the slider itself is shifted until the first $\varkappa$-curve cuts the cross-point of the two-pointers. Then, at the origin of the $\varkappa$-curves the virtual height is scaled. Only when the observed ionospheric layer is of an Epstein-, cosine- or parabolic-type will the succeeding virtual height differences be equal and result in equal values for the true height of the maximum and in equal values for the half-thickness Y_m of the layer, which is a quarter or a tenth of Y_m according to the differences of the parameter values $a - 0$ of the $\varkappa$-sheets (Fig. 4).

Thus, all the interesting data of an ionospheric layer can be obtained within 8 min. It is also possible to obtain some limited information about the ionization valley between the *F1*- and *E1*-layers from the delayed penetrating echoes. Fig. 6 shows the respective $\varkappa_D$-curves again for a cosine layer. The parameters h'_a to h'_h measure the additional delays of ordinary echoes penetrating such a cosine layer.

VOL. 16 1959



5(a)



5(b)

Fig. 5(a)–(b). Equipment for true height calculations.

VOL.
16
1959

In order to determine the true height and the half layer thickness of the above layer when two layers are present, tables have been calculated which show the contribution of a penetrated Epstein, cosine or parabolic layer to the virtual heights of the above layer. The heights of the upper layer, determined by the respective $\varkappa$-curves, must be reduced using these values. Table 1 is an example of such a table calculated for a cosine layer. α measures the ratio f_0F2/f_0F1. The

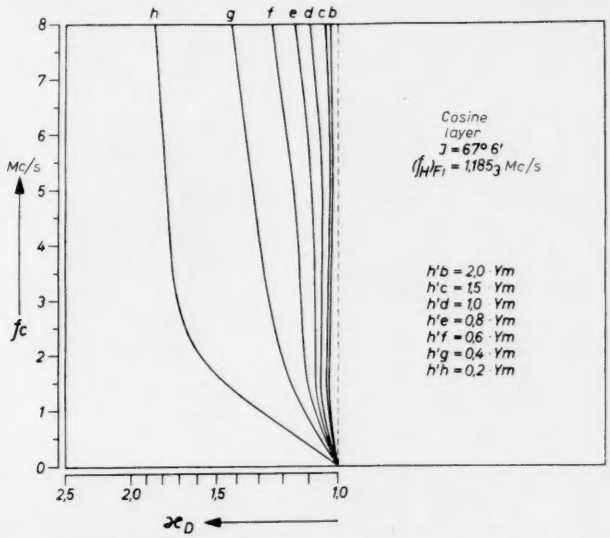


Fig. 6. Ordinary critical frequency f_c v. frequency of observation $f(\varkappa_D = f/f_c)$ for constant delays of ordinary echoes penetrating a cosine layer. Y_m measures the equivalent parabolic half layer thickness, J the angle of dip of the earth's magnetic field and f_H the gyrofrequency.

parameters a to d indicate by how much the height, obtained from the $\varkappa$-curves $a - 0$ for a single layer, is to be reduced. It is clear that when a is about 2 the dispersion of the delay time due to a lower layer is very small; therefore, the influence of the underlying layer on the proposed determination of the thickness of the higher layer is negligible, also.

1.2. *Determinations from extraordinary* h'(f)*-traces*

The method under discussion exhibits a principal and a technical difficulty. The principal one is that the ordinary and the extraordinary traces of a sweep-frequency record overlap each other over a wide frequency range. Unfortunately, the lower edge of the composite trace usually belongs to the extraordinary component. Therefore, inaccurate ordinary height readings are possible.

The technical disadvantage of the forementioned method follows from the fact that sounding equipments do not start at 0 c/s, as any solution of the respective ordinary integral equation demands, but at 0·5 or 1 Mc/s. Such a great extrapolation range from 0 to 1 Mc/s may be the cause for many errors. This disadvantage can be overcome by extending the method under discussion to the extraordinary component. That is, since the extraordinary heights are usually smaller than the ordinary heights and the mathematical beginning of the extraordinary

Table 1. Cosine layer

α \ f_0F1	1·592	1·691	1·790	1·890	1·989	2·238	2·486	2·984	3·183	3·382	3·581	3·780	3·979	4·476	4·973	5·968	6·366	curve
1·005																		
1·068	3·65	3·34	3·10	2·90	2·72	2·36	2·74	1·86	1·78	1·72	1·68	1·65	1·62	1·56	1·52	1·45	1·430	
1·131	0·815	0·805	0·786	0·810	0·803	0·771	0·776	0·763	0·765	0·768	0·772	0·777	0·783	0·800	0·822	0·869	0·885	
1·194	0·495	0·501	0·507	0·512	0·518	0·530	0·541	0·562	0·570	0·577	0·586	0·594	0·601	0·621	0·640	0·677	0·690	
1·257	0·365	0·374	0·381	0·388	0·396	0·413	0·428	0·455	0·465	0·475	0·484	0·493	0·501	0·520	0·538	0·570	0·580	*a*
1·414	0·250	0·258	0·266	0·273	0·279	0·293	0·304	0·323	0·330	0·337	0·343	0·349	0·355	0·367	0·378	0·398	0·405	
1·571	0·195	0·201	0·206	0·210	0·215	0·225	0·235	0·249	0·255	0·260	0·265	0·269	0·272	0·282	0·290	0·305	0·310	
1·885	0·130	0·136	0·142	0·146	0·151	0·160	0·166	0·176	0·180	0·182	0·183	0·185	0·186	0·189	0·191	0·194	0·195	
1·005																		
1·068																		
1·131	1·12	0·986	0·965	0·950	0·937	0·915	0·905	0·891	0·890	0·890	0·890	0·892	0·895	0·906	0·921	0·960	0·975	
1·194	0·580	0·582	0·583	0·584	0·586	0·591	0·597	0·613	0·620	0·624	0·636	0·644	0·652	0·672	0·691	0·728	0·740	
1·257	0·405	0·412	0·420	0·426	0·433	0·447	0·461	0·486	0·494	0·505	0·513	0·522	0·530	0·548	0·566	0·599	0·610	*b*
1·414	0·265	0·272	0·279	0·285	0·291	0·305	0·317	0·337	0·345	0·352	0·358	0·364	0·370	0·383	0·395	0·418	0·425	
1·571	0·205	0·211	0·216	0·220	0·225	0·235	0·244	0·260	0·265	0·272	0·276	0·281	0·285	0·295	0·305	0·320	0·325	
1·885	0·140	0·146	0·152	0·157	0·161	0·170	0·177	0·187	0·190	0·192	0·294	0·196	0·197	0·199	0·201	0·204	0·205	
1·005																		
1·068																		
1·131	2·66	2·40	2·17	2·03	1·885	1·67	1·54	1·38	1·34	1·31	1·29	1·27	1·26	1·23	1·22	1·21	1·205	
1·194	0·770	0·763	0·756	0·750	0·745	0·735	0·729	0·724	0·725	0·727	0·731	0·735	0·741	0·757	0·777	0·820	0·835	
1·257	0·485	0·490	0·496	0·500	0·506	0·517	0·526	0·547	0·555	0·562	0·570	0·577	0·585	0·603	0·62	0·657	0·670	
1·414	0·290	0·298	0·306	0·312	0·318	0·333	0·345	0·368	0·375	0·383	0·390	0·396	0·402	0·416	0·430	0·453	0·460	*c*
1·571	0·220	0·226	0·232	0·236	0·241	0·252	0·262	0·278	0·285	0·290	0·295	0·300	0·305	0·315	0·325	0·340	0·345	
1·885	0·150	0·155	0·161	0·164	0·169	0·178	0·184	0·194	0·197	0·199	0·201	0·203	0·205	0·209	0·212	0·218	0·220	
1·005																		
1·068																		
1·131																		
1·194	1·770	1·66	1·58	1·52	1·45	1·33	1·25	1·10	1·065	1·040	1·030	1·020	1·010	1·00	1·010	1·030	1·035	
1·257	0·700	0·692	0·687	0·683	0·679	0·672	0·670	0 675	0·680	0·686	0·692	0·699	0·705	0·721	0·737	0·765	0·775	*d*
1·414	0·330	0·338	0·346	0·353	0·359	0·375	0·388	0·412	0·420	0·428	0·436	0·444	0·450	0·466	0·482	0·508	0·516	
1·571	0·245	0·250	0·255	0·258	0·263	0·275	0·285	0·303	0·310	0·316	0·322	0·328	0·333	0·344	0·355	0·374	0·380	
1·885	0·160	0·166	0·171	0·175	0·180	0·188	0·196	0·202	0·210	0·213	0·216	0·218	0·221	0·226	0·230	0·237	0·240	

$f_0F2 = \alpha f_0F1$

VOL. 16 1959

trace coincides with $f = f_H$, the extraordinary trace always lies within the frequency range of observation. It can be shown mathematically that ordinary virtual heights are not always necessarily greater than the extraordinary heights, even for a single layer. This is the case only for layers with a sharp lower edge like that of a parabolic layer (Fig. 7).

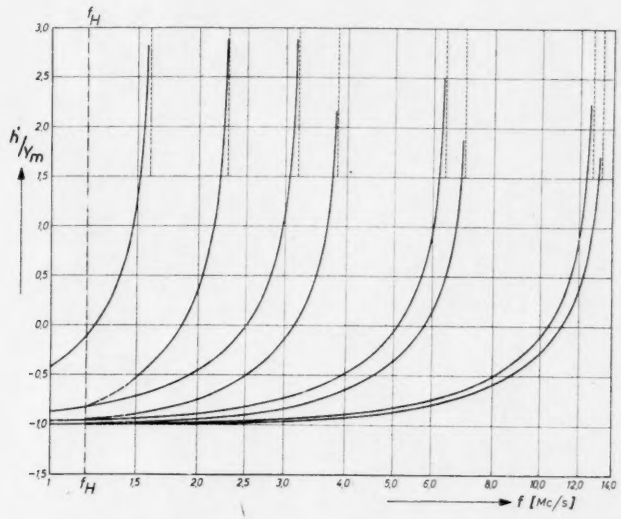


Fig. 7. Calculated $h'(f)$-curves for parabolic layer profiles of different maximum electron densities. Y_m measures the half layer thickness and f_H the gyrofrequency (1·1853 Mc/s) and J the angle of dip of the earth's magnetic field; $J = 67°6'$.

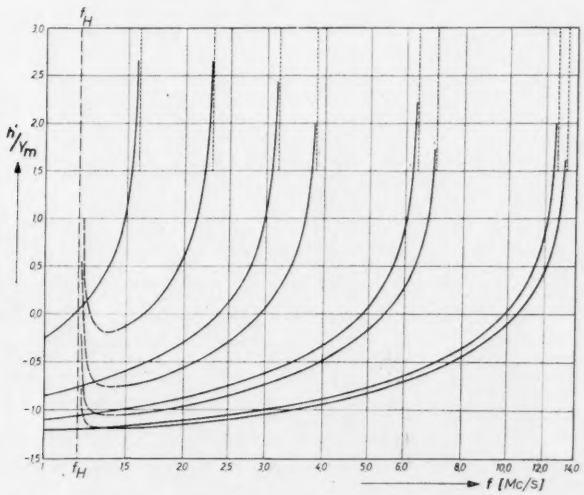


Fig. 8. Calculated $h'(f)$-curves for cosine layer profiles of different maximum electron densities. Y_m measures the equivalent parabolic half layer thickness, f_H the gyrofrequency (1·1853 Mc/s), and J the assumed angle of dip of the earth's magnetic field; $J = 67°6'$.

Fig. 7 shows the $h'(f)$-curves for a parabolic layer type calculated for four critical frequencies. The other case results where the ionospheric layer has a "tail" like that of a cosine layer (Fig. 8) or the more pronounced one of an Epstein layer (Fig. 9). By the way, the fact that the extraordinary heights are usually smaller than the ordinary heights tells us that the night-time F-layer usually has a sharp lower edge. After this digression, we return to the calculation of true heights from extraordinary virtual heights. As before, the respective $\varkappa$ values were taken from these calculated extraordinary virtual height traces. Fig. 10 shows the $\varkappa$-curves belonging to the same h'-values as before together with the ordinary curves of Fig. 4. Both samples are shifted against each other allowing a quick analysis, which is now independent of the quality of one of the traces.

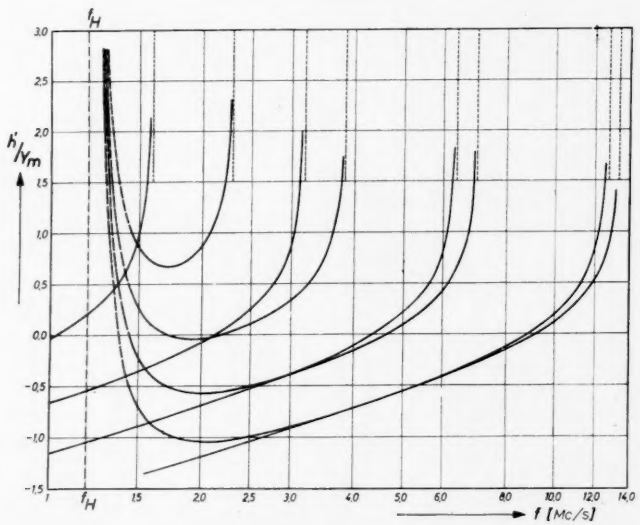


Fig. 9. Calculated $h'(f)$-curves for Epstein layer profiles of different maximum electron densities. Y_m measures the equivalent parabolic half layer thickness, f_H the gyrofrequency (1·1853 Mc/s) and J the angle of dip of the earth's magnetic field; $J = 67°6'$.

2. General Methods

This section describes general methods for routine calculations of true ionospheric heights. The calculation of the wanted extraordinary group velocities for a magnetic dip $J = 67°6'$ as a multiple U_{xL}/U_x of the extraordinary group velocities for pure longitudinal propagation ($J = 90°$), with the same value of f_H, shows that U_{xL}/U_x differs by only about ± 4 per cent from 1 (Fig. 11). This important result implies that at middle or northern stations (such as Lindau with $J = 67°6'$) Rydbeck's (1940) solution of the respective integral equation is applicable and the final accuracy is better than the measured accuracy ($\sim$2 km), provided one knows the height of the lower edge h_0 of the layer. That is the point Rydbeck had overlooked. Indeed Figs. 7–9 show that the extraordinary virtual height at

VOL. 16 1959

$f \simeq f_H$ becomes infinite if the layer has a tail. This implies that h remains undetermined. But this difficulty can be overcome with sufficient accuracy in two ways, so that RYDBECK's solution remains applicable.

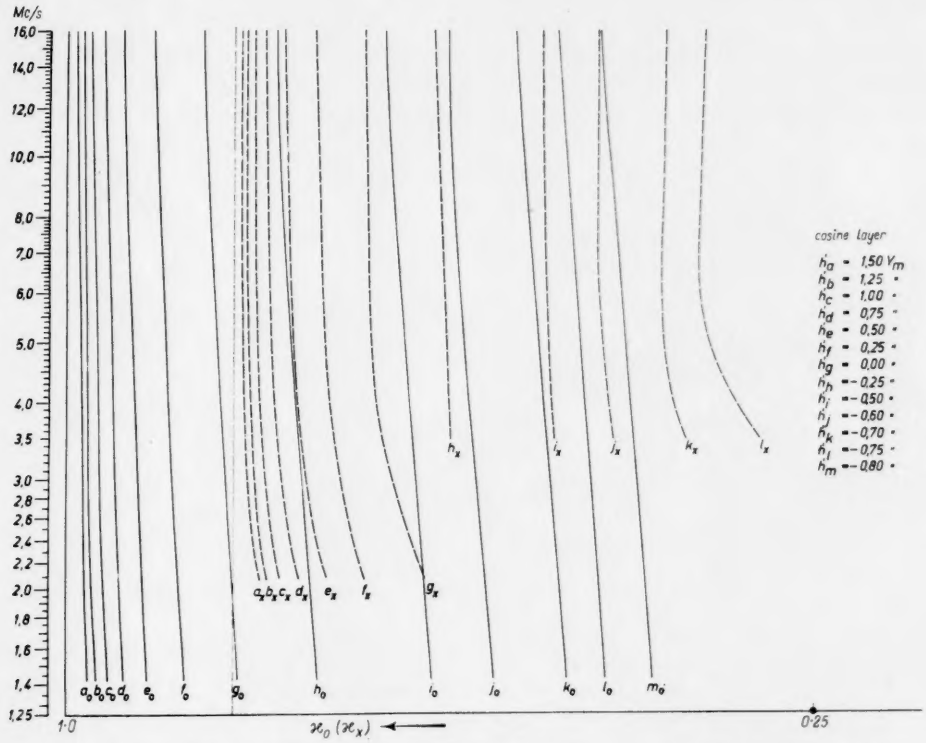


Fig. 10. Ordinary f_{c0} and extraordinary f_{cx} critical frequencies versus frequencies of observation f for constant ordinary, index "0", and extraordinary, index "x", virtual heights of a cosine electron density profile.

($\varkappa_0 = f/f_{c0}$ full curves, $\varkappa_x = f/f_{cx}$ broken curves)

Y_m measures the equivalent parabolic half layer thickness, f_H the gyrofrequency (1·1853 Mc/s) and J, the angle of dip of the earth's magnetic field, 67°6′.

(a) One must not take the lower edge of the ionosphere as the origin of the mathematical ordinate as RYDBECK has done. One should take the ground level as the origin. RYDBECK's solution can then be expressed in the following way (BECKER, 1958)

$$h(f) = h_0 J_0 + \frac{2}{\pi} \int_{f_R'=0}^{f_R'=f_R} \frac{H'(f_R')df_R'}{\sqrt{\{f_R^2 - (f'_R)^2\}}} \tag{1}$$

with

$$f_R^2 = f^2 - ff_H; \quad J_0 = \frac{4}{\pi} \int_{f_R'=0}^{f_R'=f_R} \frac{f_H^2}{4(f_R')^2 + f_H^2 + f_H\sqrt{\{4(f_R')^2 + f_H^2\}}} \times \tag{2}$$

$$\times \frac{df_R'}{\sqrt{\{f_R^2 - (f_R')^2\}}},$$

and

$$H'(f) = h'(f) - \Delta h'(f) \tag{3}$$

$$\Delta h'(f) = \frac{f_H}{2f - f_H}\ (h'(f) - \frac{1}{f}\int_{f'=f_H}^{f'=f} h'(f')\,df'). \tag{4}$$

In this, h_0 does not appear behind the integrals; it is represented by an explicit expression, it is multiplied by J_0 only, a small factor diminishing with increasing frequency f (Table 2). This shows that errors in h_0 do not introduce the same order

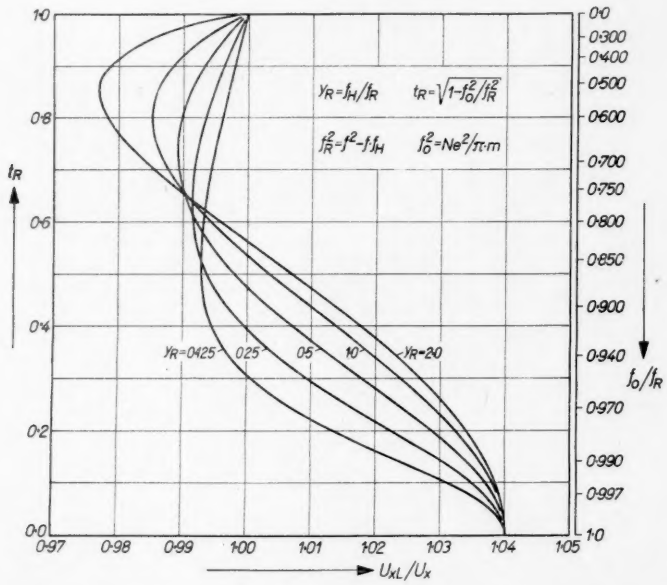


Fig. 11. Ratio U_{xL}/U_x of the extraordinary vertical group velocities for 90 (U_{xL}) and 67·1 (U_x) degrees angle of dip of the earth's magnetic field. For both group velocities a constant gyrofrequency f_H = 1·1853 Mc/s is assumed.

error in the final result. This shows the desired approximate solution of the general problem. For *routine purposes* it is suitable to make use of KELSO's integral approximation (1952) in equation (1) and SIMPSON's approximation in equation (4). In addition an extension of the optical–graphical method described in Section (1) is a great time saver. Further, a tabulation of the factors in equations

Table 2.

f	J_0	f	J_0
2·0	0·491	5·0	0·168
3·0	0·300	6·0	0·138
4·0	0·215	7·0	0·117

VOL. 16 1959

(1)–(4) is preferred to a graphical representation. The extension of the optical graphical method can be done in the following way: KELSO's approximation gives for equation (1):

$$h(f) = h_0 J_0 + \frac{1}{n} \sum_{K=1}^{K=n} H'(f_K) = h_0 J_0 + \frac{1}{n} \sum_{K=1}^{K=n} [h'(f_K) - \Delta h'(f_K)] \tag{5}$$

with

$$f_K = \frac{f_H}{2} + \sqrt{\frac{f_H^2}{4} + (f^2 - f f_H) \cos^2 \theta_K}; \quad \theta_K = \frac{2K - 1}{2n} \pi; \quad K = 1, 2, 3, \ldots n$$

Equation (5) theoretically implies that one has to take the mean of n reduced virtual heights $H'(f_K)$ and to add them to $h_0 J_0$ in order to get the desired true

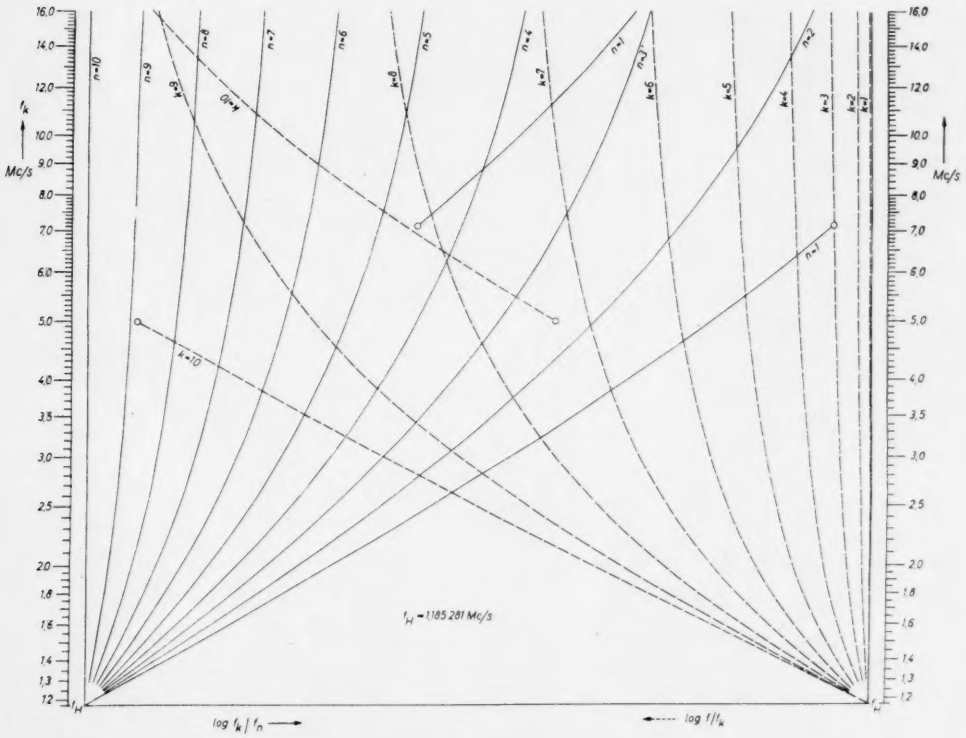


Fig. 12. Auxiliary curves for numerical integration according to Kelso (K-curves) and Simpson (n-curves).

reflection height. Practically, one has to read n virtual extraordinary heights at the respective f_K-frequencies and to determine the n corrections $\Delta h'(f_K)$.

By the optical–graphical method it is possible to find very quickly these reading frequencies f_K for any value of f. Some of the sample curves are shown by Fig. 12 which shows f vs. $\ln f/f_K$ for 10 values of K. The n-curves in Fig. 12 give 10 equally

spaced frequencies between f_H and f_K for any value of f_K for the $\Delta h'$-calculations according to SIMPSON's approximation method. The time needed for a thorough analysis of one $h'(f)$-curve according to the method in Section (2a) amounts to about 2 hr.

(b) There is another possibility for obtaining a more accurate method which

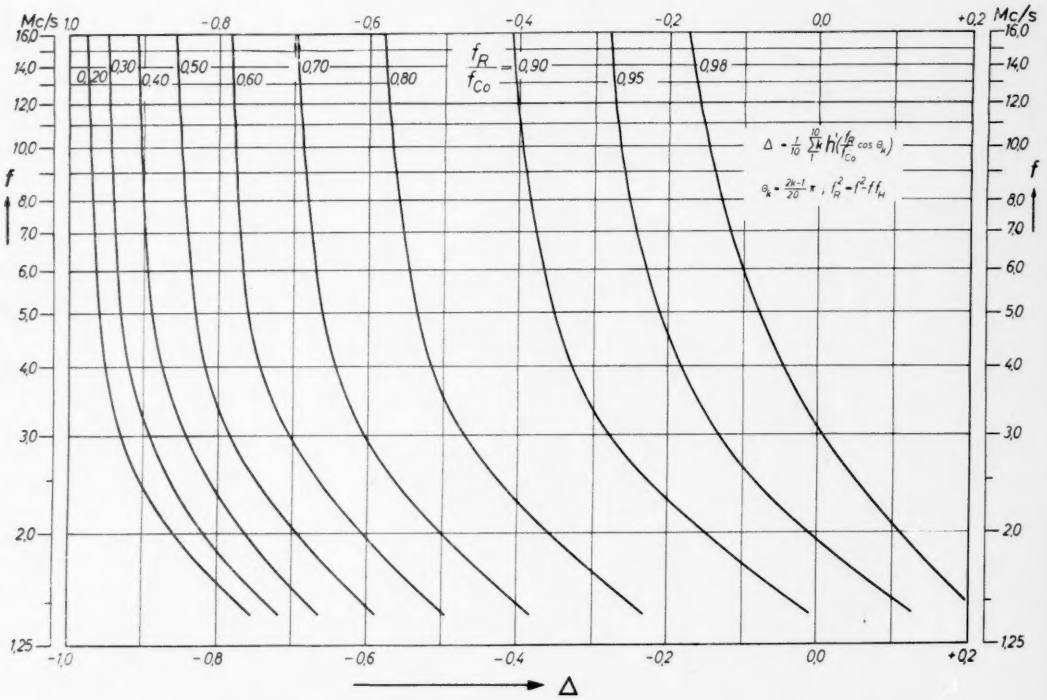


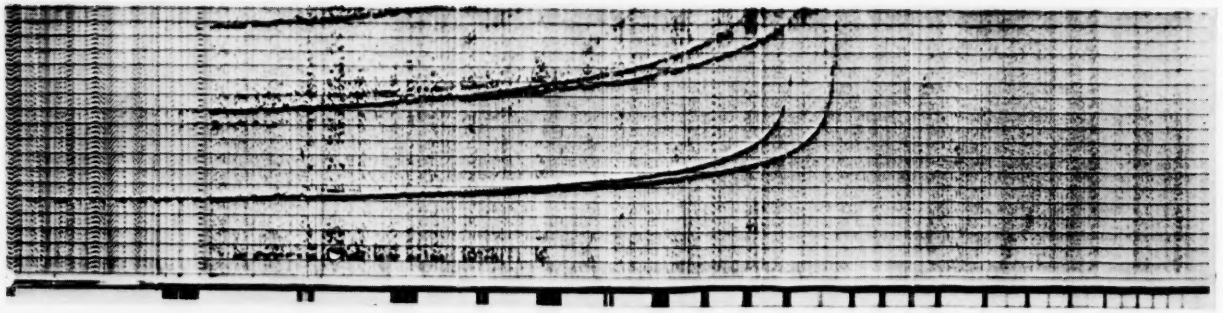
Fig. 13. Graphical representation of the second integral of equation (7) f_{c0} measures the ordinary critical frequency of the underlying parabolic electron density profile, f_R the gyrofrequency (1·1853 Mc/sec) and J the inclination of the earth's magnetic field (67·1°).

is independent of h_0, the unknown height of the lower edge of the ionosphere. We can write for equation (1):

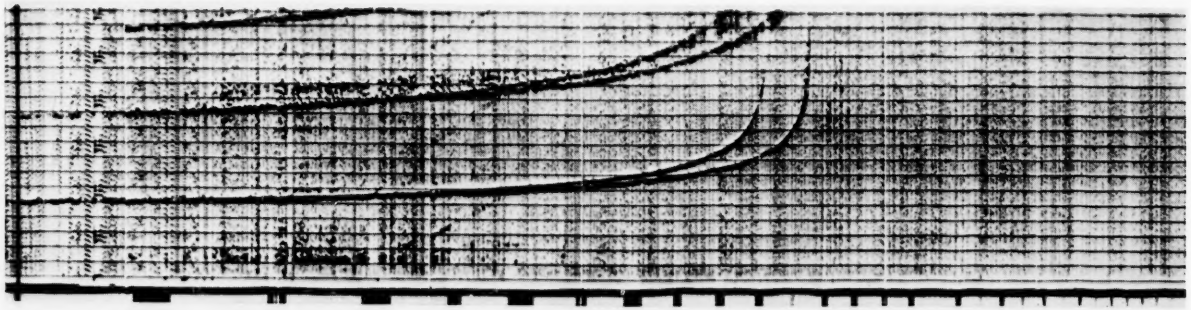
$$h(f) = h_0 J_0 + \frac{2}{\pi}\int_{f_{R'}=0}^{f_{R'}=f_R} \frac{h'(f')\,df_R'}{\sqrt{\{f_R^2 - (f_R')^2\}}} - \frac{2}{\pi}\int_{f_{R'}=0}^{f_{R'}=f_R} \frac{\Delta h'(f')df_R'}{\sqrt{\{f_R^2 - (f_R')^2\}}} \qquad (6)$$

The second integral in equation (6) is a correction to the first one. Therefore, one can say that this correction is about the same for the actual layer and the best fitting parabolic, Epstein or cosine layer determined according to the method in Section (1). The consequence is that the difference between the true reflection heights of the actual layer and this best-fitting layer is represented by the

VOL. 16 1959



14(a)



14(b)

Fig. 14. $h'(f)$-records taken at Lindau/Harz on 8 August 1957 (a) 0000 hr and (b) 0015 hr 15° EST.

VOL.
16
1959

difference of the first integrals of the appropriate equations (6). For example, we have for a parabolic layer:

$$\Delta h(f) = [h(f)]_{\substack{\text{act}\\\text{layer}}} - [h(f)]_{\substack{\text{par}\\\text{layer}}} \approx \left[\frac{2}{\pi}\int_{f_R'=0}^{f_R'=f_R} \frac{h'(f')df_R'}{\sqrt{\{f_R^2 - (f_R')^2\}}}\right]_{\substack{\text{act}\\\text{layer}}} - \left[\frac{2}{\pi}\int_{f_R'=0}^{f_R'=f_R} \frac{h'(f')df_R'}{\sqrt{\{f_R^2 - (f_R')^2\}}}\right]_{\substack{\text{par}\\\text{layer}}} \tag{7}$$

This correction Δh is to be added to the corresponding true reflection height of the best fitting (here parabolic) layer type in order to obtain the desired true reflection height. That is the principle of the correction method. The second integral is proportional to the thickness of the best fitting layer type; it is most practicable to tabulate this integral as a function of f_c and f/f_c. Fig. 13 gives a graphical representation of this data. The calculation of each integral can be done according to KELSO's method (1952) described above with the aid of the K-curves of Fig. 12 Thus, the time necessary for a calculation of ten true reflection heights of one night-time *F2*-$h'(f)$-curve amounts to about $1\frac{1}{4}$ hr.

3. Examples of True Height Calculations

Fig. 14 shows records from which the true reflection heights of Fig. 15 were calculated. The records were taken during an undisturbed night. Fig. 15 shows the difference between the actual shape and the best fitting parabolic layer. One might say that the differences are not important for, at least, statistical purposes. Fig. 16 shows the variation of these true reflection heights belonging to a series of different frequencies. Without investigating the details, one can state:

(a) Considerable height variations occur near the maximum of a layer, which makes recombination-coefficient deductions from the variation of the critical frequencies very doubtful.

(b) Periodic height variations occur with an amplitude of about 8 km and a period of about 1 hr. Assuming a travelling wave velocity of 40 m/sec one gets a wavelength of 124 km for these height variations which agrees quite well with MUNRO's results. Detailed investigations are planned.

Fig. 17 is extremely interesting. It shows that during a sudden polar commencement the *F2*-layer rises as a whole with a velocity of at least 120 km/hr or 33 m/sec. At the same time the critical frequency does not alter appreciably. In fact the layer gets a little thicker. Periodic height variations like those shown by Fig. 16 are not well marked here. The reason for this might be the $\frac{1}{2}$ hr observation intervals in comparison to the $\frac{1}{4}$ hr interval for the observations shown in Fig. 16. Special investigations on this phenomenon are being carried out at Lindau.

And now to Fig. 18 which deals with the question of an electron density minimum above the E-layer maximum. Fig. 18 shows the $h'(f)$-record analysed here.

The decrease of the virtual reflection heights of the E_s- and the F-layer demonstrates that the E-layer of Fig. 18 possesses a clear minimum. It further shows that the upper half of the E-layer was only a little thicker (0·6 km) than the lower half (Fig. 19). This result was deduced in the following way:

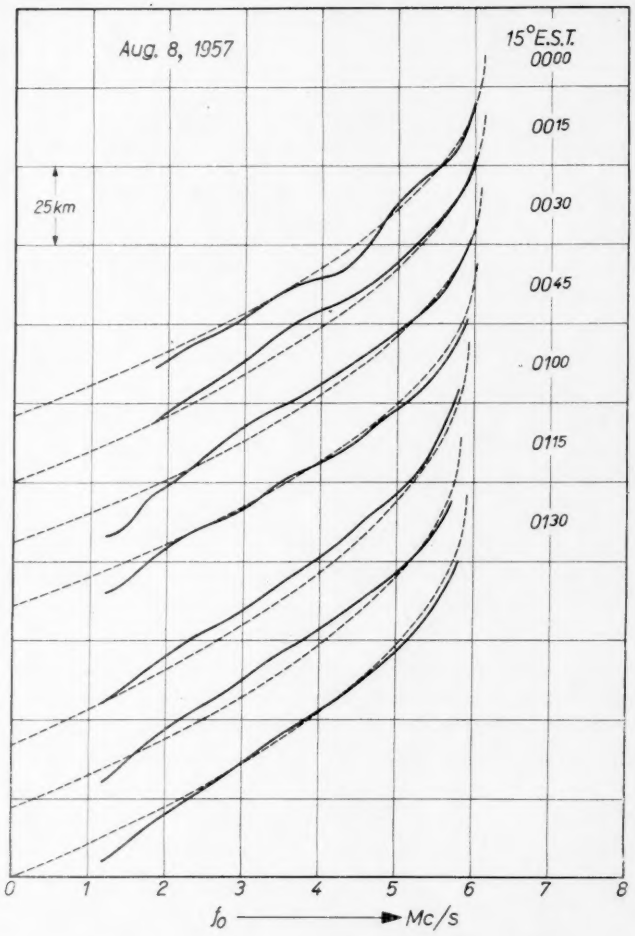


Fig. 15. $N(h)$-profiles (full curves) and the best-fitting parabolic electron density distributions. f_0 measures the plasma frequency. The profiles were deduced from $h'(f)$-records taken at Lindau/Harz on 8 August 1957 between 0000 and 0130 hours 15° EST.

First, the height and thickness of the lower half of the E-layer was determined according to the optical–graphical comparison method described in Section (1). Second, the virtual path-lengths of the echoes penetrating the E-layer determined from Table 1 were subtracted from the observed virtual heights of the E_s- and F-layer. From the decrease of these new E_s-heights, the thickness of a best fitting upper-half parabolic E-layer was determined (16·0 km). Third, the virtual paths due to this part of the E-layer were taken from Table 1 and were subtracted from the *corrected* E_s- and F-virtual heights. A constant height of the E_s-layer of 117 km, i.e. a height well above the E-layer, maximum, was obtained.

The new F-virtual heights gave a thickness of a half parabolic layer between

VOL.
16
1959

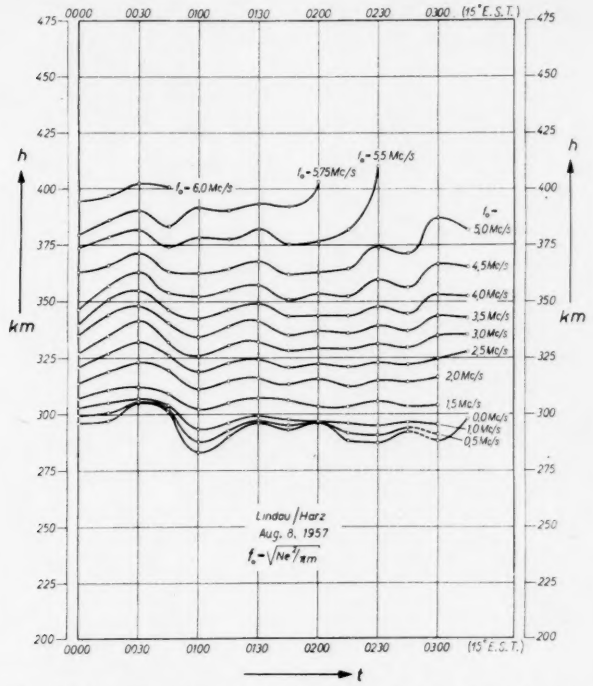


Fig. 16. Height variation of constant ionization levels deduced from the electron density profiles of Fig. 15. f_0 measures the plasma frequency.

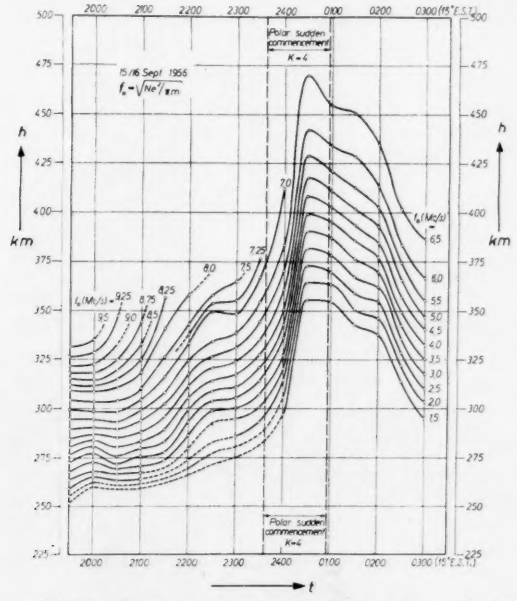


Fig. 17. Height variations of constant ionization levels above Lindau/Harz during a "Polar Sudden Commencement".

the E- and F-layer of 44 km. There is little doubt that the electron distribution between the E- and F-layers is best represented by a half layer with an inflection point at the transition to the F-layer rather than by a full layer. This can be seen in the following way:

A constant F-reflection height demands nearly equal virtual F-height steps at the selected reading frequencies of Fig. 6. A smaller step was observed for the

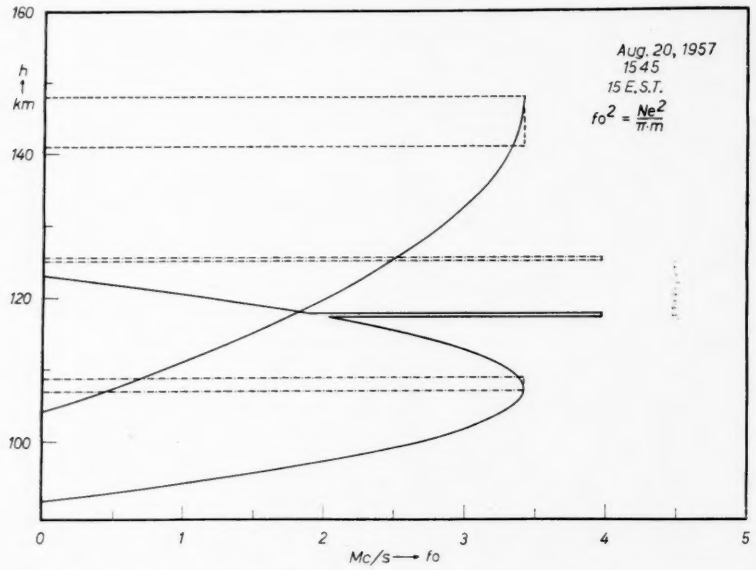


Fig. 19. Electron density profile of the E-region deduced from Fig. 18 by means of the comparison method. The full curves were obtained by assuming a parabolic profile of the upper half of the E-layer which together with the lower half causes the delay of the E_s-echoes. The broken curves show the corresponding results for rectangular profiles.

first interval, then an increase and finally a decrease in the virtual height differences. The only explanation for this observation appears to be a quick increase followed by a relatively slower increase of the true reflection height. This implies that a half parabolic layer is the most realistic result. Unfortunately, extraordinary $\varkappa$-curves similar to the ordinary curves shown by Fig. 6 are not yet calculated. Therefore, true-height calculations for such a corrected monotonous extraordinary trace have not been carried out. But this work is in preparation.

Fig. 20 indicates that the error can become significant if one assumes a continuous transition from the E- to the F-layer trace. This figure was calculated according to the more general method in Section (2a).

Conclusion

For statistical investigations of true height variations of summer night-time and winter day and night-time $F2$-layer traces, the optical–graphical comparison method in Section (1) is satisfactory. Furthermore a comparison of extraordinary

VOL.
16
1959

OL.
16
959

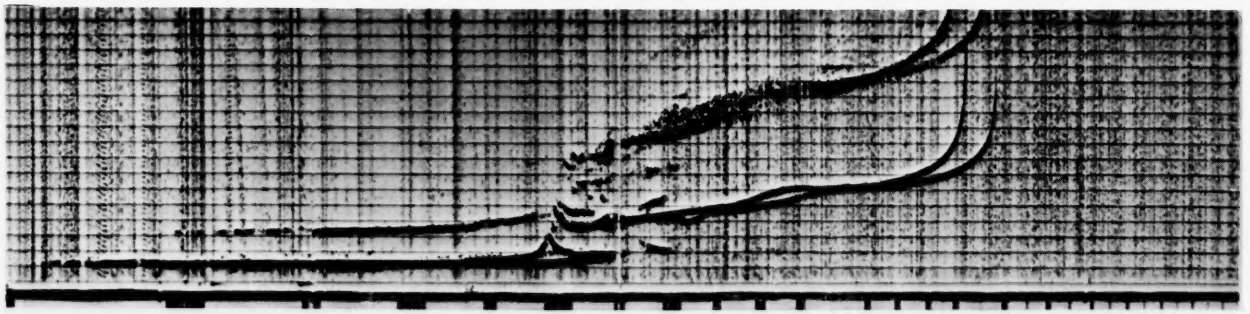


Fig. 18. $h'(f)$-record taken at Lindau/Harz on 20 August 1957, 1545 hours 15° EST.

VOL.
16
1959

traces is preferable. The time necessary for such a full comparison is about 8 min. Fine structure investigations are most quickly carried out with the more accurate method in Section (2b), the correction method. One needs $1\frac{1}{4}$ hr for the correction of ten true reflection heights obtained according to the comparison method in Section (1). The generalized method in Section (2a) is only applicable to monotonous $h'(f)$-curves; these may be given or corrected according to the method in

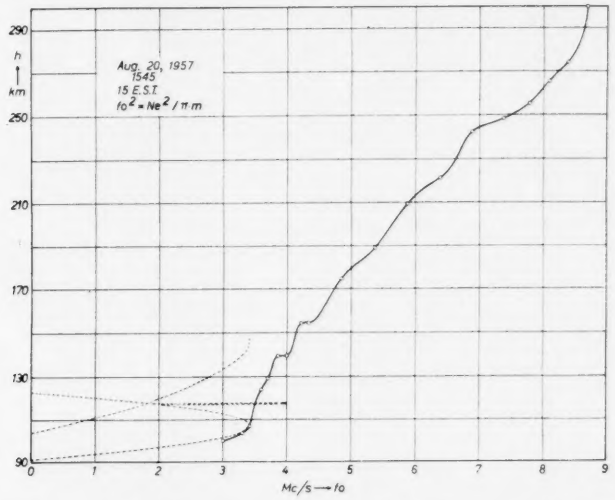


Fig. 20. Electron density profile of the ionosphere deduced from Fig. 18 by means of the general method in Section 2(b). The full curve assumes a monotonic $h'(f)$-trace; the broken curves agree with Fig. 19.

Section (1). Minimum investigations can only be carried out with the comparison method. Only the generalized method is restricted to stations with more or less pure longitudinal propagation (middle and high latitude stations). The other methods are principally free from any restrictions.

References

Becker W.	1955	*Arch. Eleptr. Übertr.* **9,** 277.
Becker W.	1956	*Arch. Eleptr. Übertr.* **10,** 207.
Becker W.	1957	*Arch. Eleptr. Übertr.* **11,** 166.
Becker W.	1958	*Arch. Eleptr. Übertr.* In press.
Budden K. G.	1954	*The Physics of the Ionosphere* p. 332. Physical Society, London.
Kelso J. M.	1952	*J. Geophys. Res.* **57,** 357.
	1954	*J. Atmosph. Terr. Phys.* **5,** 11.
King G. A. M.	1957	*J. Atmosph. Terr. Phys.* **11,** 209.
Rydbeck O. E. H.	1940	*Phil. Mag.* **30,** 282.
Shinn D. H.	1954	*The Physics of the Ionosphere* p. 402. Physical Society, London.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 84 to 92. Pergamon Press Ltd. Printed in Northern Ireland

The early morning *E2*-layer and some evidence of pre-sunrise *F*-layer "splitting"

P. BANDYOPADHYAY
Institute of Radio Physics and Electronics, University of Calcutta, Calcutta, India

(*Received* 22 *October* 1958)

Abstract—Two early morning phenomena, one relating to the *E2*-layer cusps and ridges and the other to the *F*-layer traces as observed in the *h'-f* records made at Haringhata (Calcutta), are described and discussed. The early morning *E2*-layer cusps and ridges are found to be a regular sunrise feature of the ionograms. They show marked seasonal variation of character and frequency of occurrence and are most prominent in winter. The early morning *F*-layer records show, also in winter, a peculiar kind of "splitting" which is quite distinct from regular *F1*-, *F2*-bifurcation. This *F*-layer "splitting" and its possible bearing on the *E*-layer phenomena described above are discussed. Representative ionograms of the two phenomena are given.

1. INTRODUCTION

WHEN *h'-f* records are taken during day-time they generally contain three main cusp-shaped traces due to reflections from the *E*-, *F1*- and *F2*-regions of the ionosphere. In addition to these, a number of other cusps, usually much smaller ones, often appear in the record at or near the ends of the regular *E*- and *F*-layer traces. These have been observed by many workers and variously described as cusps, ledges or ridges. Of these the cusps that appear between the *E*-layer and the *F*-layer traces particularly near sunrise are the ones with which we are concerned in this study.

VOL. 16 1959

These cusps are observed quite frequently and last for hours together. They appear at heights intermediate between those of the *E*- and the *F*-layer and are generally ascribed to an *E2*-layer. The name "sequential E_s" has also been used by some workers. They have been vaguely classified as the "morning type", the "afternoon type", etc., depending on the part of the day when they are observed. Although some work has been done on the "forenoon type" and the "afternoon type" the early morning ones do not seem to have received adequate attention. The information on these early morning cusps as is already available is summarized below.

SKINNER *et al.* (1954) have studied the frequency of occurrence of the *E2*-cusps at Ibadan at different hours of the day and found three maxima: one near sunrise, one in the forenoon and one in the afternoon hours. Their study is confined mainly to the last two. Regarding the first they have simply observed that it is intimately related to sunrise effects. RASTOGI (1956) at Ahmedabad (India) has observed that the morning cusps appear regularly with sunrise between the *E*- and the *F1*-layers and exhibit some of the properties of a separate regular layer formed by solar radiation. Mention of sunrise cusps in general has also been made by MUNRO and HEISLER (1956). They, however, regard these as due to oblique reflection from curved isoionic surface rather than being indicative of a separate layer.

It may be noted in this connexion that BHARGAVA (1952) has also reported

the occurrence of cusps near sunrise on successive days at Kodaikanal (India). But these are of a type different from those described above. They appeared at great heights (300–600 km), beyond the critical frequency of the *F*-layer (instead of at heights intermediate between those of the *E*- and *F2*-layers) and descended quickly. Further, the duration of the phenomenon was quite small.

The early morning *E2*-layer at Haringhata (Calcutta) was noted by SAHA (1956), but no detailed study was made. The present study has been undertaken primarily with a view to examining the characteristics of this early morning *E2*-layer as observed by us regularly at Haringhata during the period 1955–1957. Attempt has been made to discriminate between the two theories of its origin, namely, separate layer theory and oblique reflection theory. The characteristics studied are mainly the frequency of occurrence of the *E2*-layer and its diurnal and seasonal variations and the evidences obtained are more in favour of the former.

While studying the characteristics of the early morning *E2*-layer we came across a peculiar *F2*-layer phenomenon taking place generally before sunrise in winter. This phenomenon has certain striking features and its exact parallel does not appear to have been reported anywhere. The phenomenon, as studied from a sequence of ionograms, is characterized by the gradual development of a cusp in the otherwise normal *F*-layer trace. The *F*-layer seems to split up into two: an "upper" layer and a "lower" one. The phenomenon has been described by us as pre-sunrise "splitting" of the *F*-layer. Details of the pre-sunrise "splitting" are given. The phenomenon is discussed along two different lines. The "splitting" is first regarded as only apparent and attempt is made to attribute it to effects of travelling disturbances or of sunrise tilt in the ionosphere. The difficulties of this explanation are pointed out. The "splitting" is next regarded as real and the identity of the layers is discussed. It is found that while the "upper layer" is without difficulty identifiable as the main *F*-layer the identity of the "lower layer", cannot be definitely established. A possible bearing of the phenomenon on the formation of the early morning *E2*-layer is also discussed.

It is to be noted that studies on sunrise effects in the ionosphere have been specially recommended during the IGY. Results of the present study will, therefore, interest other workers in the field.

2. The Early Morning *E2*-layer Cusps

An examination of the *h'-f* records obtained regularly at Haringhata (22° 56′N, 88° 31′E) since December 1954 shows that near sunrise the development of the *E*-layer is generally associated with stratifications. The stratifications sometimes appear as small and rather indistinct ledges in the *E*-layer trace. They may also appear as cusps—quite large and well-defined, commonly identified as the *E2*-layer trace (URSI, 1949) and precede development of the regular *E*-layer. A typical variation of the frequency of occurrence of these cusps at sunrise in the different months of a year is shown in Fig. 1. The data given in the figure are those obtained during the year 1955. For comparison, similar curve for Ibadan for the year 1952 is also given (SKINNER *et al.* 1954). It will be seen that the Haringhata values show a sort of annual variation with very low values occurring in summer; the Ibadan values do not exhibit any such trend.

Associated with the annual variation of the percentage occurrence, the character of the cusps also exhibits a gradual change round the year. Typical development of the *E2*-layer cusps in different seasons of a year is illustrated in Figs. 2 and 3 by a number of sequences of ionograms. The first sequence (Fig. 2a) is for early winter (November–December) in which season the features of the phenomenon are quite prominent. The layer appears shortly before ground sunrise (0620) as a fairly large cusp at the low frequency end of the *F*-layer trace at about 200 km. As the morning advances, i.e. with increasing cos χ (where χ is

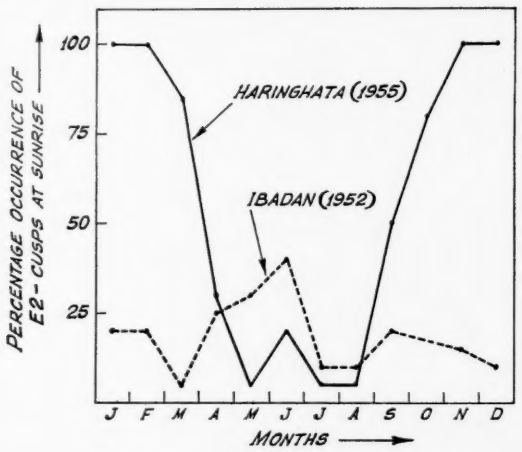


Fig. 1. Percentage occurrence of the *E2*-cusps.

the solar zenith angle) it becomes smaller and smaller in size till it forms a small ridge at the low frequency end of the *F*-layer trace. The *E*-layer trace follows in its wake (0710). The formation of the cusps exhibits the same general character during the period late winter (January–February) to vernal equinox (March–April) though the features become progressively dull (Figs. 2b and 3a). The cusps are smaller in size, show less group retardation at maximum frequency at the first appearance (a comparison between the first record of each of Figs. 2a, 2b and 3a will illustrate this) and disappear more quickly. In summer (May–August) the only indications of the *E2*-layer cusps, aupresent at all, are the indistinct ledges in the *E*-layer trace (Fig. 3b). In the if tumnal equinox months (September–October) the features begin to grow more prominent again.

The fact that at any given hour in the morning the size of the cusps grows smaller as we pass from winter to summer shows that this seasonal variation of the cusp exhibits the same kind of dependence on cos χ as the diurnal variation. This is because at any given hour of the day cos χ increases from winter to summer. A closer examination of this point shows that the fairly large early morning *E2* cusp is observable when cos χ is less than 0·1.

While the *E2*-layer first appears in the daily ionograms in the above general way throughout the year, its subsequent variation till late morning hours shows certain other types of changes besides the one illustrated in Fig. 2(a). With the advance of the morning the cusps become smaller, and more and more diffuse,

VOL. 16 1959

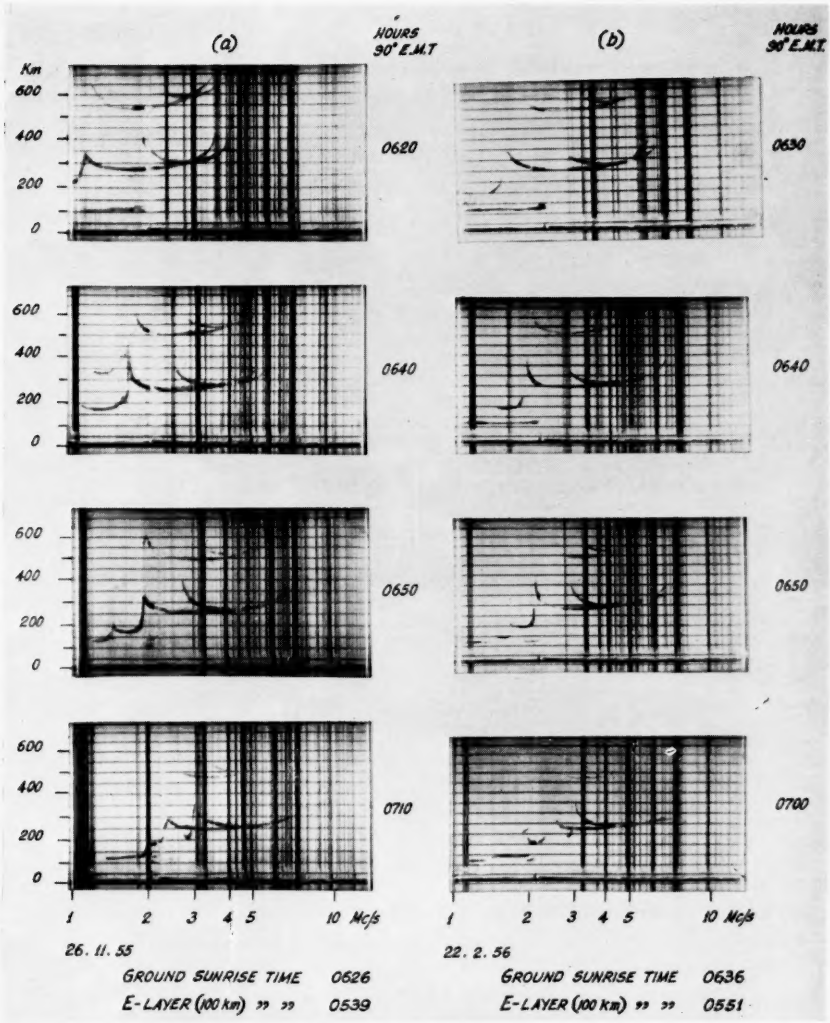


Fig. 2. Development of the early morning *E2*-layer:
(a) early winter (November–December); (b) late winter (January–February).

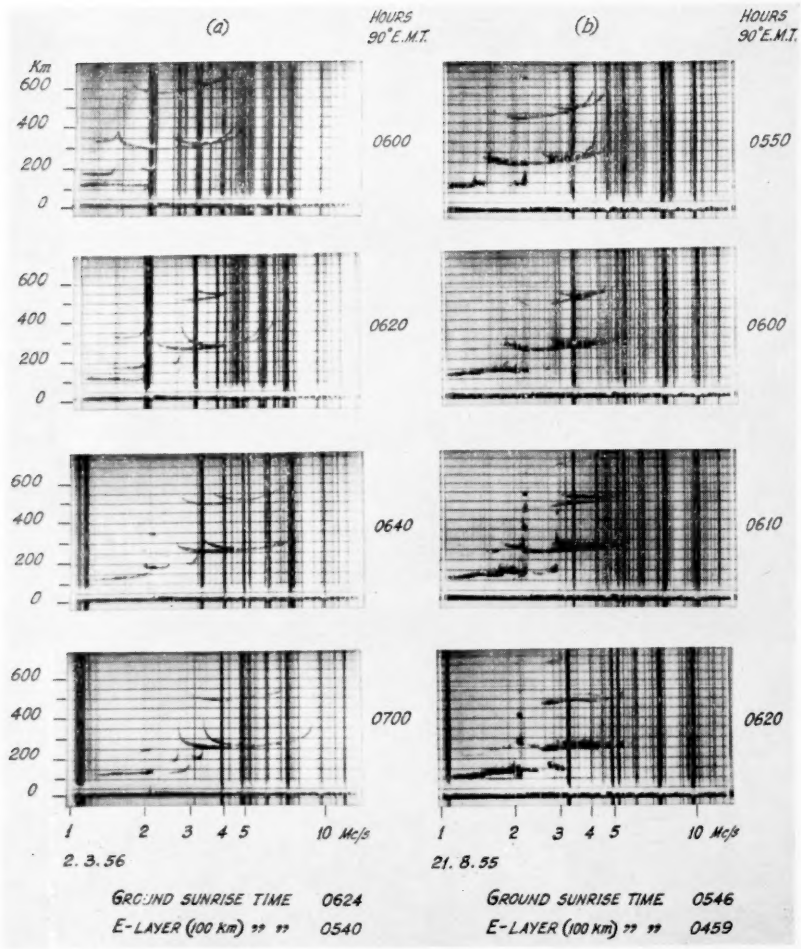


Fig. 3. Development of the early morning $E2$-layer:
(a) vernal equinox (March–April); (b) summer (May–August).

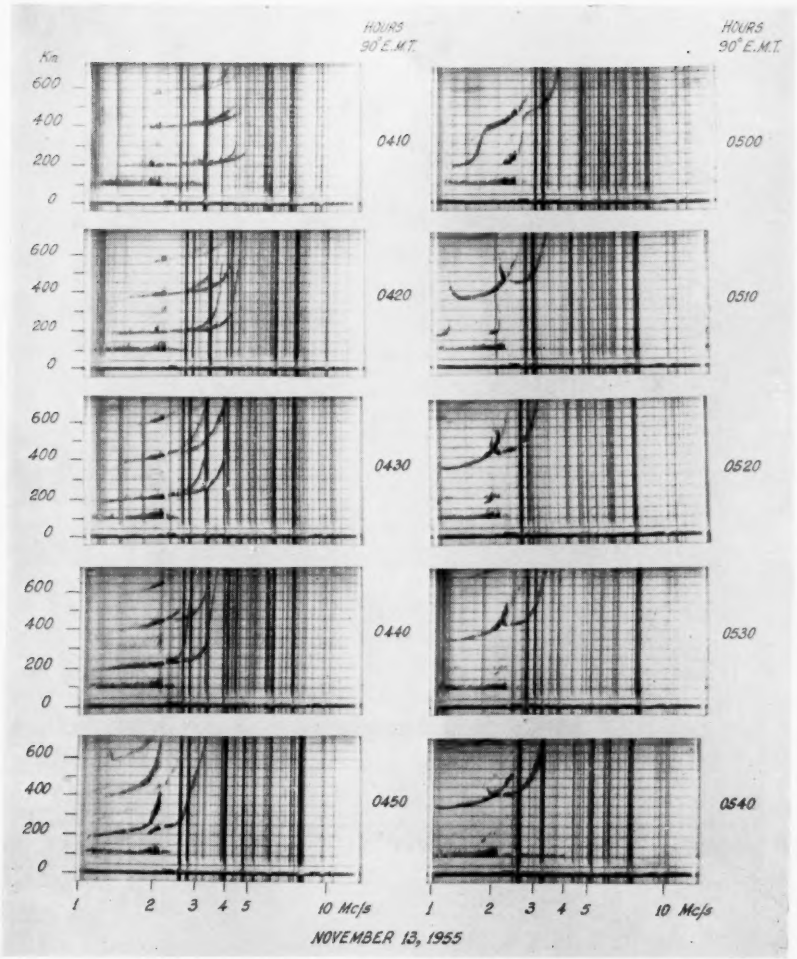


Fig. 4(a). A sequence of ionograms depicting pre-sunrise *F*-layer "splitting".

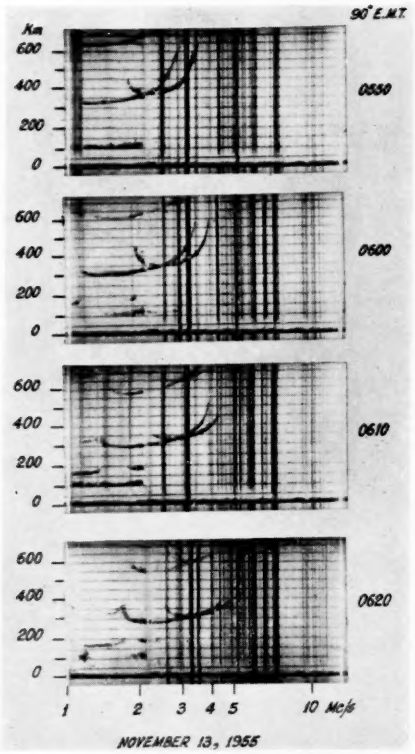


Fig. 4(b). A sequence of ionograms depicting pre-sunrise F-layer "splitting".

VOL.
16
1959

from which the *E*-layer trace takes shape. Alternatively, the cusps may gradually change into E_s-trace classified as the *c* or the *h* type in the routine reduction of ionograms. Sometimes also a third cusp is seen in addition to the two traces due to the *E2*- and the *E*-layer, but this is rare.

3. Pre-sunrise "Splitting" of the *F*-layer

In this section we shall describe an interesting phenomenon which we have come across in course of our above study of the early morning *E2*-layer. This phenomenon called here pre-sunrise "splitting" of the *F*-layer is found to occur during early morning hours mostly in the winter months (October–February). A typical sequence of ionograms presented in Fig. 4 will illustrate the phenomenon. This sequence was obtained on 13 November, 1955 from 0410–0620 hours (90°E meridian time). At 0410 hours the *F*-layer traces look quite normal. The group retardation effect near the critical frequency then gradually becomes larger, the layer becoming rather thick, its shape deviating from the nearly parabolic shape of night-time. At 0450 hours there appears a distortion notably in the *X*-trace (at 2·1 Mc/s). This distortion in the otherwise smooth traces, apparently indicative of an additional critical frequency, we call a "kink". The next record shows that the kinks, now clearly identifiable in both the ordinary and the extraordinary traces, have grown quite prominently and have moved up considerably. As a result there appears to be a "splitting" of the *F*-layer into a "lower" and an "upper" layer. At 0510 hours the "splitting" is complete. In the next three records the traces due to the "lower" layer have faded away gradually (presumably due to absorption) and in the 0550 hour record these are altogether absent. In the next record the "lower" layer traces seem to reappear but the subsequent records show that these are due to the first appearance of the *E2*-layer. The normal *E*-layer follows in the wake.

The characteristic variation of ordinary ray critical frequency of the layers referred to in the above sequence is shown in Fig. 5 as an *f*-plot together with the virtual heights of the layers (*h*-plot). The hours of sunrise at different heights have also been indicated on the abscissa.

Although the sequence of ionograms in Fig. 4 is described as illustrating typical pre-sunrise "splitting", certain minor variations of the phenomenon have been observed. These are noted mainly in the time taken by the *F*-layer to split up completely and in the sense of the apparent vertical movement of the kink. It has been found that the kink sometimes moves downwards also, and the critical frequency of the lower layer goes below the lower limit of the recorder so that the lower layer can no longer be "seen", though the sequence of changes strongly suggests its presence.

The most important characteristics of the pre-sunrise *F*-layer "splitting" phenomenon may be summarized as follows:

(1) Larger group retardation effect at the critical frequency of the *F*-layer and characteristic change of the layer shape.

(2) Appearance of a kink in the trace with an apparent vertical movement, the kink growing prominent with time.

(3) Sudden increase of the virtual height of the *F*-layer.

(4) Marked group retardation effect at the low frequency end of the traces marking the disappearance of the kink.

The features, listed above, are considerably heightened in the winter months (November–January) not remaining so prominent either in October or in February. The frequency of occurrence also passes through a similar peak during November through January, the phenomenon occurring on about 40 per cent of the days in each of these months and falling off to about 10 per cent on either side of this

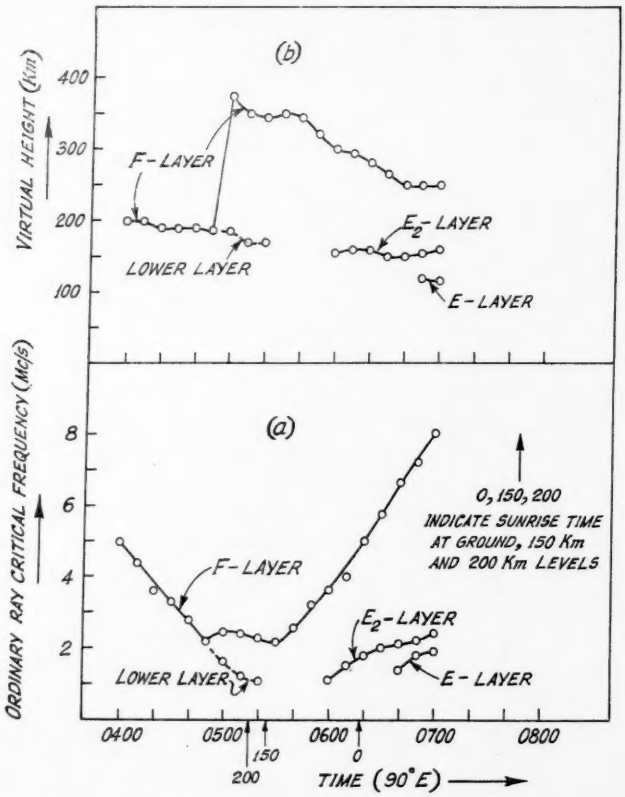


Fig. 5. Variation of virtual height and critical frequency of the layers.

period, i.e. in October and February. The phenomenon shows a tendency to recur for two or three days in succession. It occurs mostly when f_0F2 is rather low but shows no association with magnetic storms.*

4. Discussion

4.1 *Early morning* E2*-layer phenomenon*

The results of Section (2) show that the early morning *E2*-layer is a regular feature of the ionograms here, especially during the winter months. Appearing

* *Note added in proof:* We have very recently come across a paper by Mason (1958) in which he has reported formation of a "lower layer" at about 200–220 km, only in winter, which has some similarity with the one reported here. But unlike ours the "lower layer" reported by Mason was associated with small magnetic disturbances.

VOL. 16 1959

shortly before ground sunrise it shows normal layer characteristic, namely, that the critical frequency increases with decreasing solar zenith angle. The seasonal variations also exhibit this. Besides, it has now been established from rocket results that the E-layer often possesses a fine structure (LIEN *et al.*, 1954). We are, therefore, led to believe that the $E2$-cusps indicate a distinct stratification of the E-region. However, as mentioned in the introduction, MUNRO and HEISLER (1956) offer a different explanation. According to these authors the gradual variation of solar zenith angle from east to west of any place of observation, at any time,

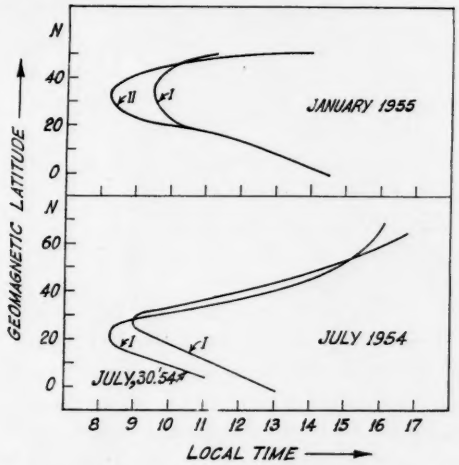


Fig. 6. First appearance of Seq. E_s at different geomagnetic latitudes (after MATSUSHITA). (I) For merger type of Seq. E_s in which the O-trace merges with the X-trace. (II) For all other types of Seq. E_s.

results in isoionic surfaces to be formed at smaller heights in the east than in the west. This results in a tilt of the isoionic surface. The sunrise cusps are then assumed to be due to non-vertical reflection from these tilted surfaces and not indicative of separate layer formation. MUNRO and HEISLER also offer the same explanation for the sunrise cusps observed by BHARGAVA (1952). However, this explanation does not seem to be applicable to the $E2$-layer at Haringhata as observed by us, because, as mentioned already, the phenomenon possesses quite different characteristics. Further, according to the tilt hypothesis, there should be $E2$-layer formation at sunset as well as at sunrise. There will also be close correspondence between the sunrise and the sunset layers since similar tilts will be present at these hours above any chosen place of observation. Observations at Haringhata show that although $E2$-layer cusps are often present near sunset these do not show any marked correspondence with the early morning ones. Sunrise tilt, therefore, does not seem to be an adequate explanation of the type of $E2$-layer under discussion.

It may not be out of place to mention here that MATSUSHITA (1955) has offered explanation of certain special type of $E2$-layer. This type is distinct from the "morning type" or the "afternoon type" and is called by him Seq. E_s. MATSUSHITA has found that the time of first appearance of Seq. E_s depends on geomagnetic latitude as shown in Fig. 6. His explanation is that this type of E_s is due to

vertical movement along the magnetic lines of force of the excess charges at the focus of the S_q electric current. The characteristic mentioned by him for recognizing Seq. E_s is that it "seems to move generally downwards and persists more than three hours". Now, the early morning *E2*-layer as observed by us often shows the same characteristics as those of MATSUSHITA's Seq. E_s, although, it appears with sunrise and not during late morning hours as is expected from Fig. 6. Further, it has been found that the early morning *E2*-layer phenomena at Haringhata often persists throughout the morning and gradually merges with the midday or the afternoon types. Thus we find that while MATSUSHITA's explanation may not be improbable the characteristics mentioned by him are inadequate for identifying his Seq. E_s with our *E2*.

4.2 *Pre-sunrise "splitting" of the* F*-layer*

Exact parallel of the phenomenon is unknown to the present author. It has, however, been found that some similarity exists between the sequence of Fig. 4 and the one described by MUNRO (1949). In the latter also is present a kink in the *F*-layer trace that moves down the curve. Although the kink looks like the appearance of another critical frequency below the *F*-layer peak, MUNRO, from simultaneous observations on a three-station system, concludes that it is not really indicative of the formation of a "lower" layer by "splitting" of the *F*-layer, but is caused by oblique reflection from undulations in the isoionic surfaces in the *F*-layer. The reflection shows itself up as an additional layer in the *h'-f* record. The undulations, according to MUNRO, resulted from travelling disturbances of compressional-type in the layer. We may seek an interpretation of the sequence of Fig. 4 also as due to an effect of some type of travelling disturbance. The type of disturbance must then possess several distinguishing features because, neither the kinks observed by MUNRO are so prominent as those of Fig. 4, nor do his published records show any clear indication of the "lower layer". Further, the Munro type disturbance occurs generally during the day, more precisely between 0700 and 1600 hours (MUNRO 1953), but the "disturbance" under study has been found regularly only near sunrise with a marked tendency to recur for a few days in succession. The observed kink being due to the Munro type of disturbance seems, therefore, very unlikely. It may also be noted in this connexion that according to RATCLIFFE (1954) large scale "ripples" generally appear on the surfaces of stratification in the *F*-layer and that these have a forward sloping wave front. As a result, the horizontal movement of the "ripples" produces an apparent downward movement of some irregularity (e.g. a kink) in the *F*-layer trace of the corresponding ionograms. If the observed kinks were due to such "ripples" they would then have always moved downward. But here the kink has also been found to move upwards on occasions.

The observed kink being essentially a sunrise phenomenon we may also try to understand the cusp shaped trace of the "lower layer" by the sunrise tilt theory (Section 4.1) of MUNRO and HEISLER (1956). MUNRO's cusps generally appear at great heights at the high frequency end of the normal *F*-layer trace and seem to descend with time to the normal height of the layer. The "lower layer" cusps as observed by us are, however, formed at the low frequency end of the *F*-layer

VOL. 16 1959

trace and *below* the normal F-layer height. Further, they often appear before sunrise in the F-layer when the tilts are not likely to be present. Hence MUNRO's tilt theory also cannot explain the observed phenomenon. Incidentally, it may be mentioned that MUNRO and HEISLER (1956) have observed disturbances during the day time reaching up to E-layer heights, but the details about these are not known.

Finally, we may regard the pre-sunrise "splitting" as real and try to identify the "upper" and the "lower" layers; the former presents no difficulty for an examination of its nature in the sequence of Fig. 4 and also the variation of its critical frequency (Fig. 5) clearly show that it is the normal F-layer. Regarding the "lower layer", however, it is difficult to say anything definitely though the following aspect is significant. In Fig. 5 if one examines the two curves depicting the variations of the critical frequencies of the "lower layer" and the $E2$-layer it will be noted that the second curve due to the $E2$-layer seems to be a continuation of the first due to the "lower layer" persisting at frequencies below the lower limit of the recorder (1 Mc/s). The "lower layer" may thus really be the $E2$-layer. The question of the frequencies of occurrence of the two phenomena—the early morning $E2$-layer formation and the pre-sunrise "splitting" of the F-layer—may be examined in this connexion. Both the phenomena attain a maximum in the same season of the year (winter). This may not be an accidental coincidence; there may be causal relationship between the two. It may be that the early morning $E2$-layer is formed on all days by the "splitting" of the F-layer due to some unknown cause. The discrepancy between the F-layer splitting being observed only on 40 per cent of the days and the early morning $E2$-layer formation on 100 per cent may be explained by the assumption that the "splitting" is of daily occurrence, but that on 60 per cent of days it had occurred below the lower limit of recorder. The phenomenon was therefore missed on these days.

However, notwithstanding these plausible evidences in favour of the kinks being indicative of the appearance of a new penetration frequency we cannot be very definite about it. This is because the kinks bear very close resemblance to those observed by MUNRO, and, these latter can be satisfactorily explained by the travelling disturbance hypothesis. We would therefore regard our interpretation—the kinks are precursors of real splitting—as tentative rather than final. The final interpretation must wait for more conclusive evidence.

Acknowledgements—The work described in the paper was carried out with financial assistance received from the Council of Scientific and Industrial Research, Government of India, and forms part of the programme of the Radio Research Committee. The author is indebted to Professor S. K. MITRA, F.R.S. for helpful discussions. He is grateful to Professor J. N. BHAR for encouragement and advice. His thanks are also due to Dr. A. K. SAHA and Dr. S. S. BARAL for suggestions and discussions.

REFERENCES

BHARGAVA B. N. 1952 *Nature, Lond.* **170,** 983.

LIEN J. R., MARCOU R. J., ULWICK J. C., AARONS J. and MCMORROW D. R. 1954 *Rocket Exploration of the Upper Atmosphere* p. 223. Pergamon Press, London

MASON R. F.	1958	*N.Z. J. Geol. Geophys.* **1,** 519.
MATSUSHITA S.,	1955	*J. Geomagn. Geoelect., Kyoto* **7,** 91.
MUNRO G. H.	1949	*Nature, Lond.* **163,** 812.
MUNRO G. H.	1953	*Proc. Roy. Soc.* A **219,** 447.
MUNRO G. H. and HEISLER L. H.	1956	*Aust. J. Phys.* **9,** 343.
RASTOGI R. G.	1956	*J. Atmosph. Terr. Phys.* **9,** 71.
RATCLIFFE J. A.	1954	*The Physics of the Ionosphere.* Physical Society, London.
SAHA A. K.	1956	D.Phil Thesis, Calcutta University.
SKINNER N. J., BROWN R. A. and WRIGHT R. W.	1954	*J. Atmosph. Terr. Phys.* **5,** 95.
U.R.S.I.	1949	*Information Bulletin on Terminology and Symbols.* Commission III, 10.

VOL.
16
1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 93 to 98. Pergamon Press Ltd. Printed in Northern Ireland

Anomalous ionospheric reflection during solar eclipses

W. L. PRICE
Radio Research Board (Sydney Laboratory), Commonwealth Scientific and Industrial Research Organization, Electrical Engineering Department, University of Sydney

Abstract—The appearance of complexities in ionosonde records has been ascribed by MUNRO to multiple reflections from inclined isoionic surfaces. Further extension of this work by MUNRO and HEISLER has suggested that irregularities appearing on ionosonde eclipse records and affecting consequent interpretations may be due to a similar cause.

In this paper a necessary relation between height and curvature of the reflecting isoionic surface to produce complexities is deduced. It is further shown by calculating a typical set of isoionic contours for solar eclipse conditions that the necessary height curvature relation is satisfied at certain regions within the eclipse zone so that complexities can occur.

1. INTRODUCTION

HORIZONTAL variation of ionospheric ionization has aroused increased interest in recent years, and records of ionosonde and fixed frequency sounding equipment ($h'f$, $h't$) have been examined for evidence of such gradients. These supply measurable information about some physical properties of the upper atmosphere; such as the large scale travelling ionospheric disturbances (T.I.D.) described by MUNRO (1949, 1950, 1953), MUNRO and HEISLER (1956), and HEISLER (1958). In recent papers MUNRO and HEISLER (1956, 1957) have shown that many of the ionospheric effects caused by solar eclipses (BEYNON and BROWN, 1956) produce irregularities in records which closely resemble those of T.I.D. and that the irregularities in both cases probably result from reflections by inclined isoionic surfaces. They point out that in all cases allowance should be made for the possibility of such tilts when analysing the records. The travelling eclipse disturbances (E.I.D.) cover a considerably larger area than T.I.D. so that inclination of the layers is in general correspondingly smaller. This paper shows how sufficient warping of the contours to account for the anomalies in records can occur.

2. OCCURRENCE IN RECORDS OF COMPLEXITIES DUE TO NON-VERTICAL REFLECTIONS

The appearance at any instant on an $h'f$ record of more than one ordinary (or extraordinary) ray trace except those due to multiple reflections will be referred to (after MUNRO) as a complexity.

Certain necessary conditions for the simultaneous reception by a station of rays from two separate parts of a reflecting layer are given in Section (2.1.1) *or* in Section (2.1.2) and in Section (2.2).

2.1.1. The layer may be a continuous surface $ACDBA$ (Fig. 1) whose slope varies but is everywhere continuous and whose normal nowhere makes an angle greater than 90° with the vertical (in practice never greater than about 10°). The station at S receives rays from areas A and B. Continuous lines such as $ACDB$ can be drawn in the surface and every small element $\boldsymbol{CD}$ ($\boldsymbol{ds}$) lies in a plane OCD normal to the surface. O is the centre of curvature of CD relative to the earth and $\boldsymbol{C'D'}$ ($\boldsymbol{ds'}$) is

the intercept on the (flat) earth by OC, OD. Then O must be sometimes above and sometimes below the earth. (This follows from the fact that $\int_A^B \boldsymbol{ds}' = 0$ and that the sense of $\boldsymbol{ds}'$ changes when O moves across the earth's boundary as ds moves continuously from A to B.) This result is equivalent to saying that there must be lines in the continuous contour whose radius of curvature R relative to the earth is less than the slant height of the layer h.

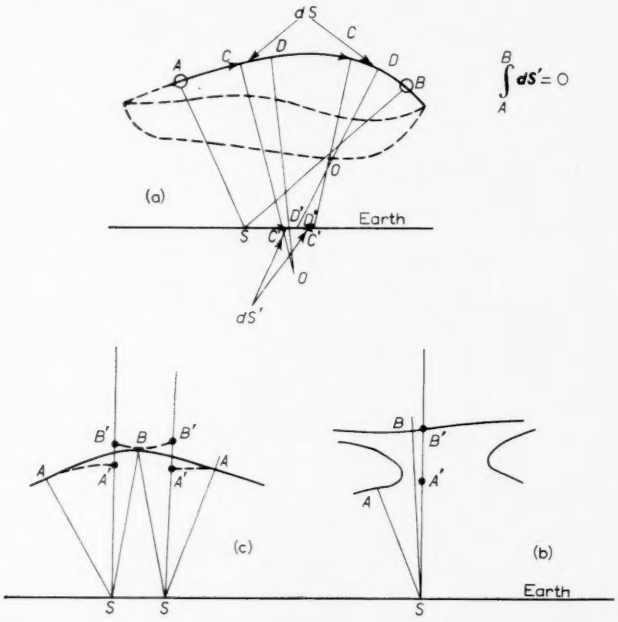


Fig. 1. Necessary conditions for the simultaneous reception by a station of rays from two separate parts of a reflecting layer.

2.1.2. The isoionic surface has an effective discontinuity, or a discontinuity of slope; then a complexity may occur as in Fig. 1(b) and Fig. 1(c).

2.2. The equipment must be able to resolve the complexity. Especially, the transmitted beam must be wide enough to fall on both reflecting surfaces and each surface must reflect enough radiation to record a trace on the film.

The optical shadow cast on the surface of the earth by a total eclipse is approximately a circle of radius 2500 km; the time from first to last contact is about 2 hr. The ionospheric region affected extends behind the shadow because of the lag due to a finite rate of recombination. This "tail" is more pronounced in the higher layers where the recombination is slower.

Consider now the maximum curvature relative to the earth, occurring when the curvature changes symmetrically throughout and uniformly in each part, being zero at the ends and maximum in the middle and having inflections at the quarter points. Suppose the distance apart of the quarter points $2x = 3000$ km, and the elevation above them of the centre $d = 20$ km. Then the radius of curvature

VOL. 16 1959

$R \simeq x^2/3d \simeq 38{,}000$ km. This is much greater than the height of the layer so that a complexity cannot occur if the curvature changes uniformly through the eclipse. There is the possibility however that some contours might change rapidly in height over a small part of the eclipse. To examine this possibility a set of contours has been calculated using probable values of recombination coefficient α and day-time (initial) electron densities N.

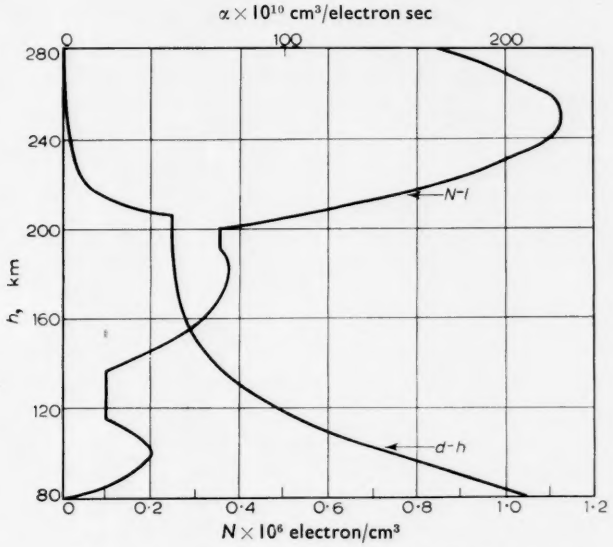


Fig. 2. Assumed variation of α and of initial N with h.

3. Calculation and Analyses of the Contours

The α–h and initial N–h curves are plotted in Fig. 2. N–h has been made parabolic for each layer (*E*, *F1*, *F2*) excepting for sections between *E* and *F1*, and *F1* and *F2* where N is assumed to be constant. It is probable that N does remain large in these "unseen" parts (Jackson, 1956); the presence of small fluctuations there would be unimportant to our main conclusions.

The actual magnitudes of h_m, N_m and the layer thickness Y_m are considered typical for day-time and mid-latitude and medium intensity distribution. No allowance is made for the zenith angle changing during the 2 hr duration of the eclipse.

It seems feasible that the electron loss function is "effectively" recombination (α) for *E* and *F1* and effectively attachment (β) above about 250 km. We have assumed that α falls by a factor of about 3·5 from h_mE to h_mF1, remains sensibly constant for 20–30 km and then falls rapidly by a factor of about 50 from there to h_mF2 at 250 km (Mitra, 1952; Szendrei and McElhinny, 1956). This variation agrees with the idea first advanced by Bradbury (1938) that *F1* and *F2* are produced by the same band of solar radiation.

The electron density N through the centre of a total eclipse was calculated

from curves of N/N_0 as a function of time for various values of a parameter $z = N_0 \alpha t_0$ given by RYDBECK and WILHELMSSON (1954) where N_0 is the initial value of ion density. It has been assumed that apparent diameters of sun (d) and moon (D) are equal and that the eclipse duration $2t_0$ is equal to 8000 sec.

The results are plotted in Fig. 3 as vertical sections of the contours of constant N, taken along the eclipse path. The maxima of electron density N_mE, N_mF1, N_mF2 are shown by broken lines. The horizontal and vertical scales are equal so that angles between intersecting lines appear undistorted. Horizontal distances are measured from O, the centre of the eclipse. These distances are contracted by the "flat earth" representation by an amount which increases with height but the angles between intersecting lines are not appreciably changed.

The contours are all horizontal of course just before and sometime after the eclipse. When the electron density is decreasing with time a given contour rises or falls according to whether it is below or above the adjacent maximum of N. The fact that α decreases with height causes the rise to occur more rapidly than the corresponding fall so that the heights of maxima of N increase while $N_{\max}$ decreases in magnitude. The slopes near the beginning and end of the eclipse are very small and unimportant and only the central part of the eclipse is shown. Here however certain contours show much steeper slopes and also discontinuities giving rise to complexities. This occurs as follows:

Where N originally has the same value above and below a maximum the two contours with this value sometimes merge and cross the maximum line. This can be seen at -286 km and $+414$ km in the E- and $F2$-layers respectively, the positions being indicated by arrows.

Again where N originally has a given value below a maximum only, then as N_m decreases, the contour with this value may cross the maximum and continue on above it, as for instance in the $F1$-layer at -50 km. It is easily seen that complexities can occur in either of these cases. For example the ($h'f$) curve for a station S at 470 km would show the ordinary ray trace like that in Fig. 4 with two branches in the $F1$-section. One branch produced by rays reflected as at A goes to infinite virtual height when critical reflection occurs at the level marked N_mF1; it then continues until infinite virtual height again occurs at h_mF2. The rays for the second branch are reflected at points such as B. The trace for this branch is a spur, ending when the reflecting area B becomes so small that the radiation reflected is insufficient to form a visible record.

Had a maximum existed between h_mF1 and h_mF2, which was originally obscured by h_mF1, the gradient in the recombination rate could result in it being revealed. The result would be to divide $F2$ into two, as shown by the dotted line, as if another layer had been formed.

The general sequence of events will be reversed when the ionization is increasing with time. It is clear that there is an asymmetry in the positions and times of occurrence of the anomalies because of the lag caused by a finite α. Two further points may be noted:

Bending of the rays due to refraction is negligible because the rays are all nearly normal to the strata. Focusing may occur at reflection because the station is sometimes near the centre of curvature of the reflecting surface so that $h = R$.

VOL.
16
1959

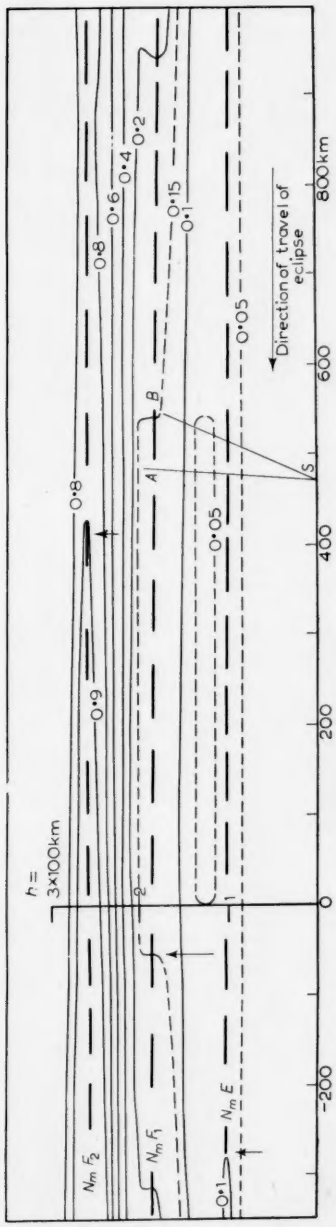


Fig. 3. Isodensity contours for electrons at the centre of a total eclipse (figures on contours are electron densities $\times 10^{-6}$).

4. CONCLUSION

Analysis of electron densities during an eclipse shows that while the slopes and curvatures of the strata are mostly very small, effective discontinuities occur and the slopes are sometimes large enough to produce complexities in the records of ionospheric sounding equipment. These complexities are due to rays reflected along paths inclined to the vertical.

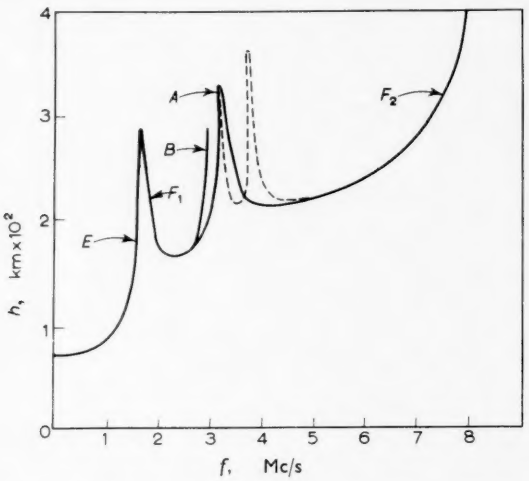


Fig. 4. Form of the ordinary ray trace which would be recorded when the station was at 470 km (Fig. 3). A complexity appears due to rays reflected from surfaces *B*. The dotted part shows the effect if a maximum, originally not seen, was between h_mF1 and h_mF2.

Acknowledgements—This work has been carried out as part of the programme of the Radio Research Board of the Commonwealth Scientific and Industrial Research Organization and is published with the permission of the Executive. The author wishes to thank Dr. G. H. MUNRO for suggestion of the problem and both him and Mr. L. H. HEISLER for helpful discussion.

REFERENCES

BEYNON W. J. G. and BROWN G. M.	1956	*Solar Eclipses and the Ionosphere.* Pergamon Press, London.
BRADBURY N. E.	1938	*Terr. Magn. Atmos. Elect.* **43,** 55.
HEISLER L. H.	1958	*Aust. J. Phys.* **11,** 79.
JACKSON J. E.	1956	*J. Geophys. Res.* **61,** 120.
MITRA S. K.	1952	*The Upper Atmosphere.* The Asiatic Society, Calcutta.
MUNRO G. H.	1949	*Nature, Lond.* **163,** 812.
MUNRO G. H.	1950	*Proc. Roy. Soc.* A **202,** 208.
MUNRO G. H.	1953	*Proc. Roy. Soc.* A **219,** 447.
MUNRO G. H. and HEISLER L. H.	1956	*Aust. J. Phys.* **9,** 343, 359.
MUNRO G. H. and HEISLER L. H.	1957	*J. Atmosph. Terr. Phys.* **12,** 57.
RYDBECK O. E. H. and WILHELMSSON H.	1954	*Chalmers Tek. Högsk. Handl.* **149.**
SZENDREI M. E. and MCELHINNY M. W.	1956	*J. Atmosph. Terr. Phys.* **9,** 118.

VOL. 16 1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 99 to 102. Pergamon Press Ltd. Printed in Northern Ireland

Sudden changes in the virtual height of radio waves reflected from the *E*-region of the ionosphere

J. D. WHITEHEAD
Otago University, Dunedin, N.Z.

Abstract—It has been observed that the virtual height of radio waves reflected from the *E*-region shows discontinuous changes. The lower height reflections take place from thin layers within the normal *E*-region. From the measurements of the associated amplitude changes it has been possible to deduce that the electronic collision frequency at 110 ± 5 km is (1·3 ± 0·1) × 10^4 per sec and the semi-thickness of the thin layers is 4 ± 2 km.

INTRODUCTION

THE virtual height of reflection and the amplitude of the echo of 2 Mc/s waves reflected at vertical incidence from the *E*-region of the ionosphere were measured at Cambridge, England, for over 100 days during 1952–1953. It was observed that the virtual height showed sudden discontinuous changes of from 5 to 30 km and

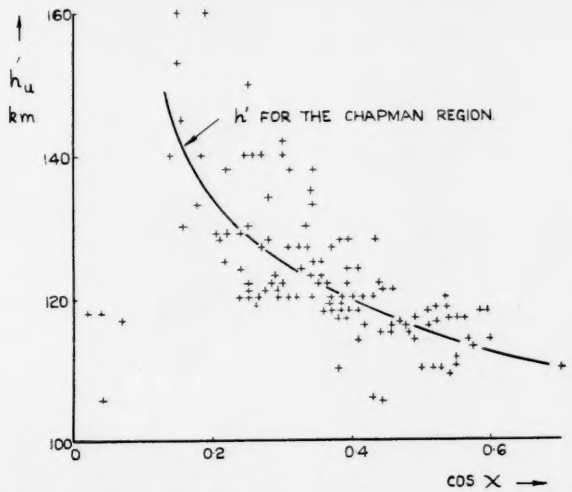


Fig. 1. A plot of the virtual height of the upper level reflection against the cosine of the sun's zenith angle.

that, at the same time, the mean amplitude of the echo, averaged over a period of 5 min, changed by a factor averaging about 2, and increased when the virtual height dropped and decreased when it rose. Over a hundred of these changes have been observed: they were more frequent in winter than summer.

The upper level reflection usually took place from a Chapman region. This may be seen by plotting the virtual height of the upper level reflection against the cosine of sun's zenith angle χ, as in Fig. 1. From the published values of critical frequency, it was calculated that when the sun's radiation was incident vertically on the ionosphere, the critical frequency of this Chapman region was 3·4 Mc/s: the corresponding height of maximum electron density and scale height were adjusted to fit the experimental points plotted in Fig. 1. This height was found

to be 110 ± 5 km and the scale height 10 ± 3 km. The solid curve shows the virtual height for 2 Mc/s waves for this Chapman region, plotted against cos χ.

The absence of group retardation during the changes in virtual height indicates that the lower level reflection took place from a thin layer. The few points lying well away from the curve on Fig. 1 probably represent jumps from one thin layer to another.

The purpose of this paper is to show how values of the collision frequency of electrons and thickness of the lower layer may be calculated from measurements of the virtual heights and ratio of the amplitudes of the echoes from the upper and lower levels.

THE METHOD OF CALCULATION

When a radio wave is reflected at vertical incidence from a Chapman region, most of the absorption occurs within the top-most few hundred meters of the trajectory (WHITEHEAD, 1956). This absorption is then given by APPLETON's (1928) well-known expression, namely

$$(h' - h)\nu/c$$

where ν = collision frequency near the actual reflection level
c = velocity of light
and h' and h are the virtual and phase heights.
The phase height is, to within 1 km, equal to the actual height of reflection.

This formula is now applied to the reflection from both the upper and lower levels.

Let h'_u, h'_L be the virtual heights for the upper and lower level reflections and h_u, h_L be the corresponding actual heights.

Let ν_0 be the collision frequency at the height h_0 of maximum electron density of the Chapman region when the sun's rays are incident vertically on the ionosphere (i.e. at 110 ± 5 km), and let H be the scale height (i.e. 10 ± 3 km). Then the collision frequency at the actual height of reflection is given by

$$\nu = \nu_0 \exp\left(\frac{h_0 - h_u}{H}\right) \text{ for the upper level reflection}$$

$$\text{and } \nu = \nu_0 \exp\left(\frac{h_0 - h_L}{H}\right) \text{ for the lower level reflection}$$

Using APPLETON's expression the absorption near the upper reflection level is:

$$\frac{\nu_0}{c}(h'_u - h_u) \exp\left(\frac{h_0 - h_u}{H}\right)$$

and that near the lower reflection level is:

$$\frac{\nu_0}{c}(h'_L - h_L) \exp\left(\frac{h_0 - h_L}{H}\right)$$

The absorption in the lower ionosphere is the same for both reflected rays and therefore the logarithm of the ratio of the two voltage reflection coefficients is given by the difference between the above expressions.

VOL. 16 1959

WHITEHEAD (1958) has shown that at high latitudes the amplitude of the echo is more nearly inversely proportional to the actual height of reflection than to the virtual height. Thus, if there was no absorption, the ratio of the amplitude of the echo from the lower level to that from the upper level would be given by $R_A = h_u/h_L$

Taking the absorption into account, we have then that

$$\ln R_A = \frac{\nu_0}{c}(h'_u - h_u)\exp\left(\frac{h_0 - h_u}{H}\right) - \\ - \frac{\nu_0}{c}(h'_L - h_L)\exp\left(\frac{h_0 - h_L}{H}\right) + \ln h_u/h_L$$

Now $\ln R_A$, h'_u, h'_L, h_0, H and c are known and h_u is calculated from the known electron density profile of the Chapman region. The third term on the right-hand side is always small and is calculated assuming that h_L is 3 km less than h'_L. The second term on the right-hand side may be written

$$\frac{\nu_0}{c}(h'_L - h_L)\exp\left(\frac{h'_L - h_L}{H}\right).\exp\left(\frac{h_0 - h_L}{H}\right)$$

Thus for each observation of a sudden change in virtual height we arrive at an equation involving only the two functions

$$\nu_0 \text{ and } \nu_0(h'_L - h_L)\exp\left(\frac{h'_L - h_L}{H}\right)$$

We assume that $h'_L - h_L$ is the same for each observation. This implies that the lower layers have constant thickness although they may occur at different heights; this would be the case if they were formed by incoming particles all of which had the same initial velocity but which were incident on the ionosphere at various angles.

Thus values of ν_0 and $h_L - h_L$ may be found from any two observations. Solutions to pairs of equations were found graphically. In practice virtual height changes involving nearly equal values of h'_u and h'_L were grouped together and weighted accordingly.

RESULTS OF THE CALCULATIONS

The median values of the two functions gave

$$\nu_0 = (1{\cdot}3 \pm 0{\cdot}1) \times 10^4 \text{ per sec (at a height of } 110 \pm 5 \text{ km)}$$

and

$$h'_L - h_L = 2{\cdot}3 \pm 1 \text{ km}$$

Provided the reflection from the thin layer occurs well away from the maximum electron density of the layer, $h'_L - h_L$ is slightly more than one half of the semi-thickness of the thin layer, assuming it to be parabolic. Thus the semi-thickness of the layer is 4 ± 2 km—rather thicker than the value JACKSON and SEDDON (1958) found for sporadic-E from rocket measurements.

The value for ν_0 is in reasonable agreement with the value found by NICOLET (1953) who gives a value of $\nu = 2{\cdot}7 \times 10^4$ per sec at 110 km and about $1{\cdot}5 \times 10^4$ per sec at 115 km.

CONCLUSION

Measurement of the changes in group height and amplitude of echoes from region-E of the ionosphere show that:

(1) The echo having the longer group path is generally reflected from a Chapman region having a scale height of 10 ± 3 km.

(2) The echo having the shorter group path is reflected from a layer of semi-thickness 4 ± 2 km (about one-fifth of the thickness of the Chapman region).

(3) From the changes in amplitude it has been shown that the electronic collision frequency at a height of 110 ± 5 km is $(1{\cdot}3 \pm 0{\cdot}1) \times 10^4$ per sec.

It is felt that further measurements of this type will lead to more accurate estimates of the parameters involved and to an understanding of the mechanism of the formation of the fine structure of E-region.

Acknowledgement—I wish to record my thanks to the staff of the Radio Section of the Cavendish Laboratory, Cambridge, under whose guidance I performed the experimental work.

REFERENCES

APPLETON E. V.	1928	*Int. Union Sci. Rad. Tel.* **1,** Part I.
JACKSON J. E. and SEDDON J. C.	1958	*J. Geophys. Res.* **63,** 197.
NICOLET M.	1953	*J. Atmosph. Terr. Phys.* **3,** 200.
WHITEHEAD J. D.	1956	*J. Atmosph. Terr. Phys.* **9,** 276.
WHITEHEAD J. D.	1958	*J. Atmosph. Terr. Phys.* **12,** 150.

VOL. 16 1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 103 to 105. Pergamon Press Ltd. Printed in Northern Ireland

The effect of the *F1*-layer on the calculation of the height of the *F2*-layer

M. D. VICKERS
Official Communication from DSIR Radio Research Station

(*Received* 20 *November* 1958)

Abstract—An approximate relationship between the estimated true height of the peak of the *F2*-layer and that given by assuming a single layer having a parabolic electron-density distribution is derived and compared with experimental data.

1. THEORY

FOR many years tabulated values of estimates of the true height of the *F2*-layer have been issued by organizations which make ionospheric measurements at vertical incidence. Two of the techniques used for producing the estimates are described by BOOKER and SEATON (1940) and APPLETON and BEYNON (1940). In these methods the *F2*-layer is assumed to be a single layer having a parabolic distribution of electron density with height, all ionization below this layer being neglected. The group retardation of a radio wave reflected in the *F2*-layer will, however, be increased as it passes through the lower layers; hence the virtual height at any given frequency will be greater than it would be in the absence of the lower layers.

This increase in virtual height is neglected in the methods referred to above for estimating the true height of the *F2*-layer peak and it leads to a positive error, Δh, which will obviously increase as the ratio $x = f_0F1/f_0F2$ approaches unity. It is possible to make a very approximate estimate of the error, provided certain assumptions are made. It is assumed that: (a) any departure from a parabolic electron density distribution for the *F2*-layer can be neglected; (b) the contribution to the group retardation, arising from sources other than the *F1*-layer, is negligible. Given these assumptions, the expression for Δh will depend on the way in which the electron density, N, varies with height in the *F1*-layer and on the semi-thickness, y_m, of the layer. Three such density distributions are considered:

(a) N does not vary with height

$$\Delta h_1 = \{(1 - x^2)^{-1/2} - 1\}y_m$$

(b) N has a parabolic distribution

$$\Delta h_2 = \{(1/x) \coth^{-1} (1/x) - 1\}y_m$$

(c) N has a linear distribution

$$\Delta h_3 = [2(1/x^2)\{1 - (1 - x^2)^{1/2}\} - 1]y_m$$

2. COMPARISON WITH MEASUREMENTS

The distribution of electron density with height has been calculated by THOMAS, HASELGROVE and ROBBINS (1957) who used a method which makes full

allowance for the *E*- and *F1*-layers. From these authors' tables, the estimated true height, h_t, of the peak of the *F2*-layer can be obtained. The bulletins of ionospheric data for Slough contain values of the height, h_m, of the *F2*-layer estimated using the Appleton–Beynon method which ignores the retardation in the lower layers. The monthly mean values, excluding the five International Disturbed Days, of $(h_m - h_t)$ for each hour have been plotted against the corresponding values of three functions representing $\Delta h/y_m$. The correlation coefficients of the plotted points did not differ significantly in the three cases, but this result was to

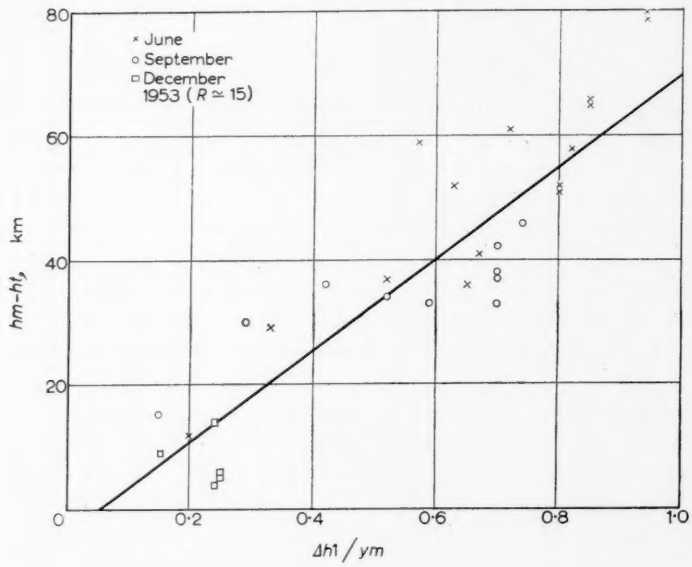


Fig. 1.

be expected as the shapes of the functions differ appreciably only when $x \simeq 1$. For a given value of x, $\Delta h_1 > \Delta h_2 > \Delta h_3$ and, since the calculated values of Δh_1 were nearest to the observed values of $(h_m - h_t)$, Δh_1 appears to be the best approximation.

Fig. 1 shows $(h_m - h_t)$ plotted against Δh_1 for a low level of solar activity ($R \simeq 15$) for June, September and December 1953. Fig. 2 contains similar data for $R \simeq 80$ for January, July and September 1950. In both cases the correlation coefficient is approximately 0·88 which indicates that the assumptions made in deriving the expression for Δh_1 are reasonable. Best-fit lines have been added in both diagrams.

The difference in slope of the best-fit lines for 1950 and 1953 means that the thickness of the *F1*-layer was not the same for the 2 years. Further, the fact that neither of the lines goes through the origin indicates that there are factors, other than the *F1*-layer, which affect the excess group retardation of the radio waves reflected from the *F2*-layer.

VOL. 16 1959

3. Use of the Diagrams

Figs. 1 and 2 are only two examples of the approximate relationship that exists between $(h_m - h_t)$ and x. Similar relations for other stations could be produced very quickly, if required, as soon as more values of h_t become available.

The method, although crude in approach, enables some idea of the height of the *F2*-layer peak to be obtained, given only the values of h_m (or the M.U.F. factor) for the *F2*-layer which are published by many organizations making ionospheric measurements.

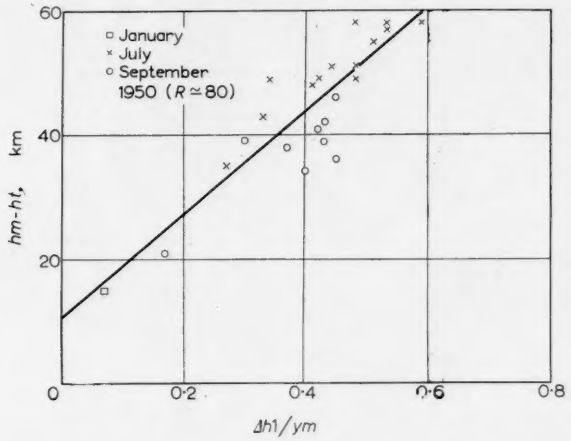


Fig. 2.

4. Conclusions

The height of the *F2*-layer, as determined by one of the usual techniques, is too high because no allowance is made for the group retardation in the *F1*-layer. An approximate relation has been derived between the height errors due to this cause and the ratio f_0F1/f_0F2. The use of this relation allows corrections to be applied to tables of height data obtained by present standard methods.

Acknowledgement—The work described above was carried out as part of the programme of the Radio Research Board. This paper is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research.

References

Appleton E. V. and Beynon W. J. G.	1940	*Proc. Phys. Soc.* **52,** 402.
Booker H. G. and Seaton S. L.	1940	*Phys. Rev.* **57,** 87.
Thomas J. O., Haselgrove J. and Robbins A. R.	1957	*Tables of Ionospheric Electron Density, Slough, England.* Series B, No. 1. Radio Group, Cavendish Laboratory, Cambridge.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 106 to 114. Pergamon Press Ltd. Printed in Northern Ireland

Variations in the geomagnetic field at Ibadan, Nigeria—I Solar variations

C. A. Onwumechilli and N. S. Alexander
Physics Department, University College, Ibadan, Nigeria

Abstract—Magnetic records obtained at the recently established magnetic observatory at Ibadan in equatorial Africa for the period November 1955 to June 1957 are analysed harmonically for solar variations. The results, given in the form of harmonic dials, show that the amplitude of H is very large (similar to that observed at Huancayo), and the amplitude of Z is larger by a factor of 3 than any previously reported. These results are explained as arising from the equatorial electrojet centred on the magnetic equator, which is about $2\frac{1}{2}°$ N of Ibadan. The variation in D at Ibadan is small, but has its phase completely reversed between the June and December solstices. The D-variation shows no obvious effect of the electrojet.

1. Introduction

A magnetic observatory, using a set of normal La Cour variometers kindly supplied on loan by the International Union of Geodesy and Geophysics, was established at the University College, Ibadan, Nigeria in 1955. The co-ordinates of the observatory are 7°26.6′N; 3°54′E; geomagnetic co-ordinates 10°39′N; 74°40′E. The mean values of the components of the earth's field are: $H = 32100\,\gamma$, $Z = 3400\,\gamma$, $D = 8°20'$W. Magnetic inclination, $-5{\cdot}85°$ approximately. Absolute values have been determined by a C.I.W. earth-inductor magnetometer kindly supplied on loan by the Carnegie Institution, Department of Terrestrial Magnetism, Washington, D.C., and by a set of three Q.H.M. and a B.M.Z.

VOL. 16 1959

The analysis of magnetic data from equatorial regions is of considerable interest because of the relatively small number of observatories operating in low latitudes (and especially in low magnetic latitudes). This applies particularly to equatorial Africa for which, so far as is known, no analysis has previously been published.

A further point of interest is that, in the longitude of Ibadan, there is a particularly wide separation of the geographical, geomagnetic and magnetic equators. This is of importance in the study of the equatorial electrojet (Alexander and Onwumechilli, 1957; Onwumechilli, 1958).

2. Analysis

Daily records of horizontal force (H), vertical force (Z) and declination (D) obtained at Ibadan on undisturbed days ($C_p < 1{\cdot}2$) during the period of 20 months (November 1955 to June 1957) have been analysed harmonically for solar variations. The results are shown in Table 1 for the following seasons: December solstice (denoted by D and comprising November, December, January and February); equinox (denoted by E and comprising March, April, September and October); and June solstice (denoted by J and comprising May, June, July and August); and for all months denoted by A. The analysis has been done month by month and the seasonal means are corrected for secular variation, smoothing and to the local time of the observatory.

In order to estimate the errors involved in the analysis, the probable error ellipses (BARTELS, 1932) have been computed for each season. The "mean error ellipse", demarcating the area within which the true mean for an infinitely large sample may be expected to lie has been obtained in the normal way by dividing the axes of the probable error ellipse by the square root of the number of months

Table 1. Showing the analysis of solar variation at Ibadan

Seasonal group of months	C_1 (γ)	ϕ_1	C_2 (γ)	ϕ_2	C_3 (γ)	ϕ_3	C_4 (γ)	ϕ_4
Horizontal force								
D	38·876	266°18′	16·897	112°58′	8·035	341°52′	3·022	196°36′
E	48·142	267 26	21·935	113 04	10·731	341 26	4·901	208 25
J	36·592	264 21	18·987	105 20	8·460	316 14	2·514	180 22
A	40·960	266 06	19·024	112 51	8·703	334 21	3·392	198 10
Vertical force (*upwards*)								
D	18·779	279°06′	11·500	121°16′	4·422	345°22′	2·231	232°15′
E	22·825	272 48	14·864	105 09	5·210	334 42	4·115	232 40
J	14·942	258 31	12·824	80 36	5·847	286 18	2·221	171 00
A	18·662	271 54	12·400	103 40	4·575	311 14	2·514	219 02
Declination (*west*)								
	(min)		(min)		(min)		(min)	
D	0·316′	340°39′	0·700′	287°12′	0·379′	127°35′	0·160′	05°07′
E	0·425	251 42	0·201	92 51	0·238	214 21	0·151	11 09
J	1·100	231 41	0·977	95 10	0·451	314 11	0·380	217 37
A	0·490	249 18	0·171	79 48	0·065	233 46	0·064	260 30

giving the mean. (For a mean to be significant, the "mean error ellipse" must not enclose the origin and in fact should be reasonably far removed from it.)

The monthly harmonic coefficients are plotted in the harmonic dials (Figs. 2, 3 and 4). Although all the mean error ellipses have been calculated, only those for the annual means are shown for H and for Z, since their significance is scarcely in doubt. For D, however, the mean error ellipses are shown for the solstitial as well as the annual means.

To illustrate the general form of the diurnal variation, and its seasonal variation, the mean hourly values, for the five international quietest days in each month are shown in Fig. 1.

During the period covered by the analysis the Zurich relative sunspot numbers (provisional) had the following seasonal means: December solstice, 127·1; equinox, 149·2; June solstice, 153·8; all months 141·7. The monthly means varied from 70·5 in January 1956 to 205·6 in June 1957.

3. Solar Daily Variation in H

The seasonal variation in H at Ibadan shows itself mainly as a change in the amplitude of the diurnal variation. The mean daily range in each of the seasonal groups D, E and J is shown in Fig. 1. The range is largest at the Equinox (121·5 γ).

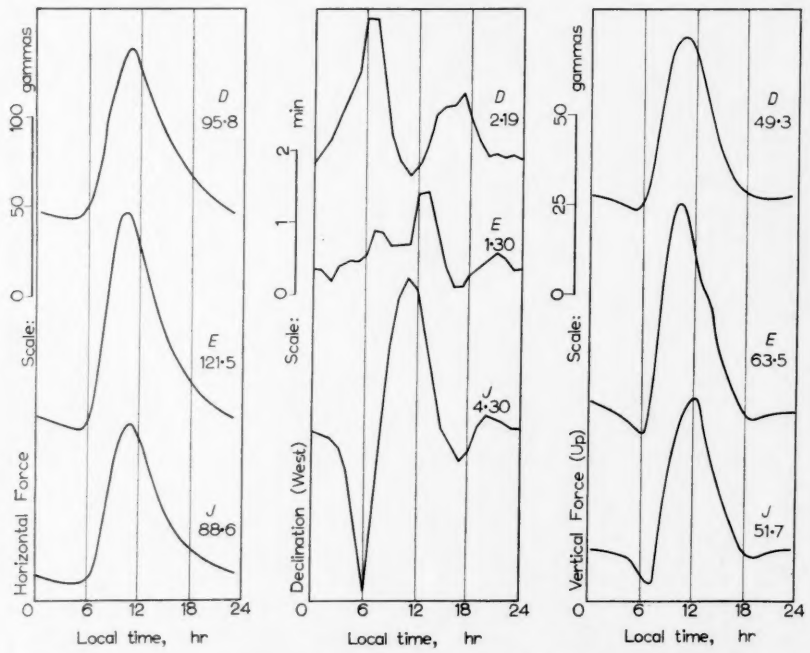


Fig. 1. Showing seasonal means of the solar daily variation of magnetic elements on international quiet days at Ibadan.

The mean range in the whole period is 103 γ on international quiet days and 100 γ on undisturbed days.

The results of the analysis for all seasons and all components are statistically significant. The figure shows that there is no significant seasonal variation in phase.

The amplitude of variation of S in H at Ibadan is much higher than the amplitude expected purely on the basis of its geographic latitude. Considering the four other tropical observatories given in Table 2, it may be seen that the amplitude of variation of H at Ibadan is about double the amplitudes at Bombay and Singapore (Sdhmidt, 1926) but is comparable to those of Huancayo (Johnston and McNish, 1930) and Kodaikanal (Thiruvegadathan, 1954). The anolamously high amplitude of H at Huancayo has been attributed to the effect of a concentration of a narrow band of current which is now called the equatorial electrojet (Chapman,

VOL. 16 1959

1951). The nearness of Ibadan to the electrojet (about 240 km) is the most likely cause of the large amplitude. It has been shown (ONWUMECHILLI, 1958) that at Ibadan, the electrojet effect on H and the "normal" variation that would occur at Ibadan in the absence of the jet are about equal. This means that without the

Table 2. Showing the analysis of solar variation at some other observatories

Observatory	C_1 (γ)	ϕ_1	C_2 (γ)	ϕ_2	C_3 (γ)	ϕ_3	C_4 (γ)	ϕ_4
North or horizontal force								
Ibadan	41·0	266°	19·0	113°	8·7	334°	3·4	198°
Huancayo	53·2	282	28·6	106	12·5	304	12·5	163
Kodaikanal	38·8	277	22·5	103	9·5	308	1·7	150
Singapore	21·5	285	9·6	112	4·3	322	—	—
Bombay	18·6	280	9·8	114	3·5	316	—	—
Vertical force								
Ibadan	18·7	272°	12·4	104°	4·6	311°	2·5	219°
Singapore	6·9	278	3·9	93	1·7	315	—	—
Bombay	5·3	84	4·5	309	4·2	148	—	—
Capetown	6·1	272	7·3	35	6·0	214	—	—
Kakioka	5·7	—	4·1	—	3·3	—	1·2	—
East force or declination in force units								
Ibadan J	10·3	232°	9·1	95°	4·1	314°	3·6	217°
Ibadan D	2·9	341	6·5	237	3·5	127	1·5	5
Singapore J	4·4	359	5·7	230	3·0	107	—	—
Singapore D	10·5	189	8·0	15	3·3	233	—	—
Bombay J	10·4	31	10·9	255	8·1	94	—	—
Bombay D	1·3	184	1·3	113	1·9	51	—	—
Capetown J	4·4	77	4·6	304	4·2	139	—	—
Capetown D	13·4	153	9·0	15	5·5	213	—	—

electrojet, the variation of H at Ibadan would have been comparable to those at Singapore and Bombay, as would be expected.

Another important feature of H-variation at Ibadan is that its first harmonic does not attain its maximum value until about local noon. For the 20 months analysed, the maximum in all seasons occurred shortly after noon (Table 1);

elsewhere, the time of maximum is nearer 1100 hours LT (cf. Table 2). It is considered that this is an aspect of the slow decline in the afternoon shown by the magnetic of H-variation taken in Nigeria south of the magnetic equator (ONWUMECHILLI, 1958). The slow decline and the delayed time of maximum have been explained as a result of the combination of the electrojet effect with a "normal" variation whose phase is slightly different.

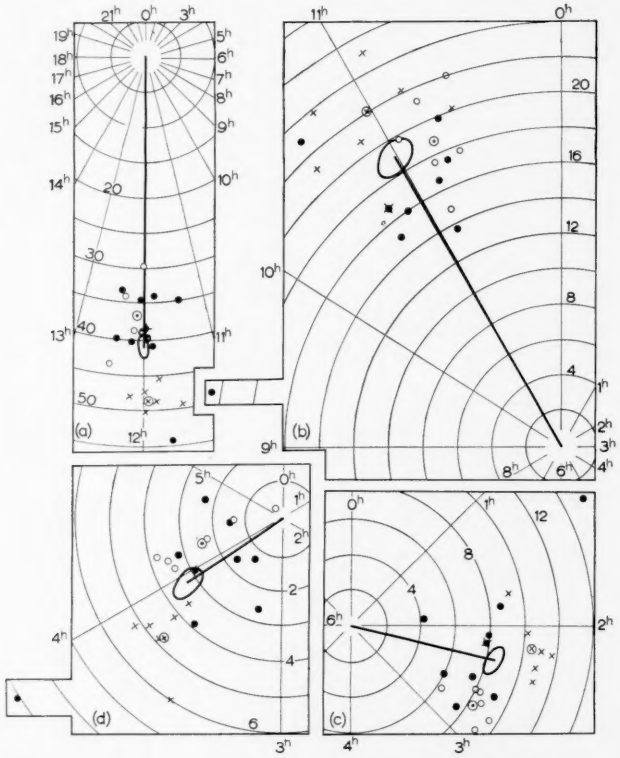


Fig. 2. Showing harmonic dials for the solar daily variation of H on undisturbed days at Ibadan: ● December solstice; × equinoxes; ○ June solstice; ⊕ mean December solstice; ⊗ mean equinoxes; ⊙ mean June solstice; (a) first component; (b) second component; (c) third and (d) fourth component.

4. SOLAR DAILY VARIATION IN Z

At Ibadan, Z is directed upwards and increases in upward Z have been recorded and analysed as positive. In Fig. 1 it is seen that the shape of the daily variation curve is much the same in the three seasons but the amplitude of variation is largest in the equinox. In the whole period covered by the analysis, the mean daily range on international quiet days is 57 γ and on undisturbed days 56 γ.

The results of the analysis (Table 1 and Fig. 3) are statistically significant for all components and in all seasons. The amplitude of S in Z-variation is anomalously large at Ibadan; being about three times larger than that so far recorded at any other observatory (cf. Table 2; in this table the values for Singapore, Bombay and

VOL. 16 1959

Capetown are from SCHMIDT (1926) and those for Kakoika are from IMAMITI, 1949). The amplitude of Z-variation at Ibadan is so large that it is comparable with that of H-variation at tropical observatories like Singapore and Bombay.

There is hardly any doubt that the Z-variation at Ibadan is largely controlled by the electrojet. It has been shown that the Z-effect of the electrojet reaches a

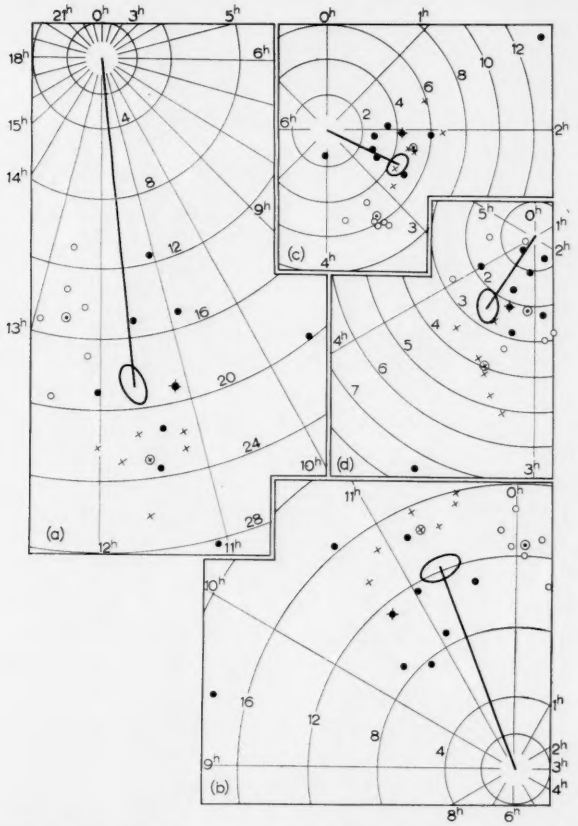


Fig. 3. Showing harmonic dials for the solar daily variation of Z on undisturbed days at Ibadan: ● December solstice; ⊗ mean December solstice; × equinoxes; ⊗ mean equinoxes; ○ June solstice; ⊙ mean June solstice; (a) first component; (b) second component; (c) third component; (d) fourth component; radii of circles are in γ.

maximum near the latitude of Ibadan (ONWUMECHILLI, 1958). Separating the observed Z-variation into the electrojet effect and the normal variation (without the electrojet), it has also been estimated that at Ibadan, the latter has an amplitude of 4·0 γ for its first harmonic component. This is of the same order as the amplitude of the same harmonic at the other observatories shown in Table 2.

The seasonal change of the amplitude of S in Z-variation at Ibadan cannot be explained by the S-current system based on the analysis of CHAPMAN (1919), according to which the northern current focus is at about 30°N latitude in June, at 40°N at the equinox and in a higher latitude still in December. Ibadan (7·45°N)

is therefore nearest to a focus in June and should consequently have its greatest Z-variation at the June solstice. The variation should be nearly zero about October when Ibadan is midway between the two oppositely directed current systems. These are not confirmed by the present analysis. Instead, it is found that Z-variation is greatest at the equinoxes and nearly equal at the June and December solstices. The most likely explanation is that at least in the longitude of Ibadan, the electrojet which controls the variation is more intense at the Equinox than at the solstices. It is also possible that the jet moves north and south according to seasons.

In the important harmonic components (first, second and third) the Z-variation reaches its maximum about $1\frac{1}{2}$ hr earlier at the December than at the June solstice (Fig. 3). Some phase difference between the normal Z-variation at these seasons is expected and to what extent this might have been influenced by the jet is not yet known.

5. Solar Daily Variation in D

At Ibadan, declination is to the West and increases in westerly declination have been recorded and analysed as positive. The seasonal variation of D at Ibadan is interesting (Fig. 1 in which the seasonal mean daily ranges in minutes are inserted), in that the form of the D-variation curve in December is an inversion of its form in June. The change of form obviously affects the daily range. On international quiet days the range in June is 4·4′; and subsequently, month after month, the depth of the minimum and the height of the maximum of the variation curve decrease. By October, the reduction is almost complete and the daily range is greatly reduced. In this transitional month of October, some days show one form of D and other days show the opposite, and this results in a mean range (1·0′) smaller than the mean of either group taken alone. After October the form, steadily changes into the December type with a mean range of 3·2′ in December. During March and April the form is again changing and the ranges are consequently reduced. The ranges of the June type are regarded as positive.

The results of the analysis of S in D are given in Table 1. On account of the phase reversal, the equinoctial and annual means are small and relatively unimportant. On the statistical test, they are hardly significant. All the results are shown in Fig. 4 in which the mean error ellipses for the solstitial and annual results are given.

The most outstanding feature of the figure is the large phase difference (almost 180°) between the June and December solstices in all the four components. The second harmonic is as large if not larger than the first as the shape of the variation curve suggests. On the whole the amplitude of D-variation at Ibadan is not abnormally large (Table 2), unlike those of H and Z.

The S-current system (Chapman and Bartels, 1940, p. 229) may be used to explain the most important feature of D-variation at Ibadan—the phase reversal between June and December. In June, Ibadan is under the northern current sheet and its D-variation is of the northern type. As the sun moves from the Tropic of Cancer towards Capricorn, the current system shifts northwards. About October, Ibadan is between the northern and southern current sheets and its

VOL. 16 1959

D-variation becomes equatorial in character. The amplitude on any day is small. It appears likely that, owing to small movements in latitude of the current system as a whole, the daily variation in the transition months when Ibadan is nearly equidistant from the two systems may on some days be of the northern and on others of the southern type. By December Ibadan is completely under the southern current sheet and the *D*-variation changes completely into the southern type. It is also seen that since Ibadan gets nearer the northern current focus in June than

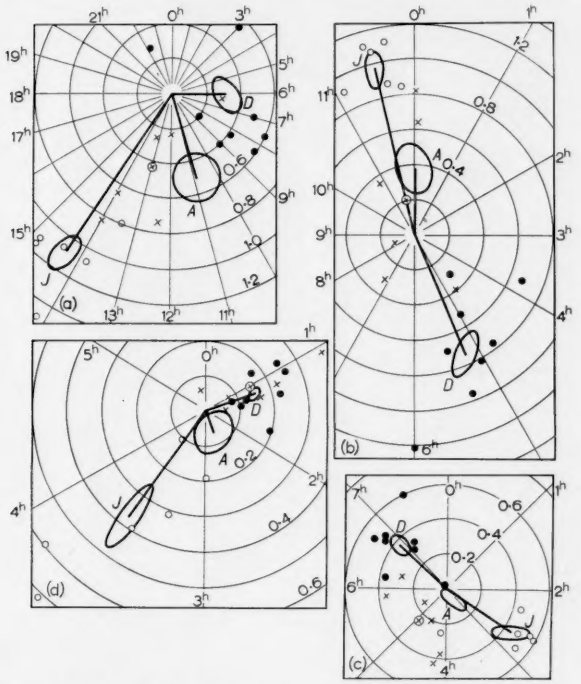


Fig. 4. Showing harmonic dials for the solar daily variation of declination (west) on undisturbed days at Ibadan: (a) first component; (b) second component; (c) third component; (d) fourth component; Radii of circles are in minutes of arc. ● December solstice; × equinoxes; ○ June solstice; ⊗ mean equinoxes; ellipses centred on the means of December solstice, all months, and June solstice are marked *D*, *A* and *J* respectively.

it does to the southern focus in December, the amplitude of variation should be greater in June, as is in fact found.

Thus the seemingly anomalous variation of *D* at Ibadan may be completely explained using the *S*-current system. The current system is based on the analysis of records from twenty-one observatories distributed all over the world (CHAPMAN, 1919). The *D*-variation at Ibadan is therefore in conformity with the world pattern of *D*-variation. It is significant that, unlike *H* and *Z* it shows no obvious effect of the equatorial electrojet.

Acknowledgements—We are grateful to the International Union of Geodesy and Geophysics for the loan of a set of La Cour variometers and to the Carnegie

Institution of Washington for the loan of a C.I.W. earth–inductor magnetometer for absolute measurements.

References

Alexander N. S. and Onwumechilli C. A.	1957	*Nature, Lond.* **180,** 191.
Bartels J.	1932	*Terr. Magn. Atmos. Elect.* **37,** 291.
Chapman S.	1919	*Phil. Trans.* A **218,** 1.
Chapman S.	1951	*Proc. Roy. Soc.* B **64,** 833.
Chapman S. and Bartels J.	1940	*Geomagnetism,* (1st. Ed.) Oxford University Press.
Imamiti S.	1949	*J. Geomagn. Geoelect., Kyoto* **1,** 12.
Johnston H. F. and McNish A. G.	1930	*Geomagnetism,* p. 242. Oxford University Press.
Onwumechilli C. A.	1958	*J. Atmosph. Terr. Phys.* In press.
Schmidt A.	1926	*Arch. Erdmagn.* Bd. I, Heft 1–4. Potsdam.
Thiruvegadathan A.	1954	*Indian J. Met. Geophys.* **5,** 267.

VOL.
16
1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 115 to 123. Pergamon Press Ltd. Printed in Northern Ireland

Variations in the geomagnetic field at Ibadan, Nigeria—II
Lunar and luni–solar variations in H and Z

C. A. Onwumechilli and N. S. Alexander
Physics Department, University College, Ibadan, Nigeria

(*Received* 20 *November* 1958)

Abstract—Magnetic records obtained from November 1955 to June 1957 at Ibadan, Nigeria, have been analysed harmonically for lunar and luni–solar variations. The results are statistically significant. The L-variations in H and Z are abnormally large at Ibadan. The L-variation in H is comparable with that of Huancayo and about three times greater than at other observatories of comparable geographic latitude. The L-variation in Z is larger than any previously reported. These results are explained as the effect of the equatorial electrojet. The ratio of S to L in H is small at Ibadan as at Huancayo. S/L in Z is equally small at Ibadan (Ibadan happens to be near the latitude where the Z effect of the electrojet is a maximum (Onwumechilli, 1959). The Chapman(1913) expression for L is fully confirmed in all aspects.

1. Introduction

The magnetic records referred to in Part I of this paper have been analysed for lunar variations (L), using the "fixed-age" method (Chapman and Bartels, 1940, p. 224). The results are given in three seasons—December solstice (D), equinoxes (E), June solstice (J) and all months (A). The data are analysed month by month to provide a test of the phase law of L. For the luni–solar variation, the same data are analysed in eight phases of the mean moon (new moon, first-eighth, first quarter, third eighth, full moon, fifth eighth, last quarter and seventh-eighth). The 3 days centred on the exact date of the lunar phase (e.g. full moon) are grouped and analysed together into Fourier components. The amplitudes and phase angles for each Fourier component are plotted against lunar age (Figs. 2 and 3).

Detailed statistical significance tests are made in order to verify the prediction of the phase law that only the second component should be significant when L is determined for a whole lunation. The probable error ellipses (Bartels, 1932) are used to test the normality of the distribution of the cloud of points formed by the monthly determinations. In order to test the statistical significance of a mean, the axes of the probable error ellipse are divided by the square root of the number of determinations forming the mean. The "mean error ellipse" so obtained demarcates the area within which the true mean for an infinitely large sample may be expected to lie. The simplest test is that, for a mean to be statistically significant, the mean error ellipse should not enclose the origin. A more rigorous test analogous to that used in scalar statistics has also been applied; namely that the mean should be at least three times the two-dimensional standard deviation; or three times the semi-major axis of the ellipse, whichever appears the more appropriate to the distribution of the cloud of points.

It is found that in all four harmonic components the probable error ellipses include approximately half the points and exclude half. It is therefore legitimate to apply the results of two-dimensional statistics of Bartels (1932) to the cloud

of points. The "mean error ellipses" for the third and fourth components include the origin showing that they are clearly not significant. The ellipse for the first component is barely clear of the origin but the ratios of the amplitude to the two-dimensional standard deviation and the semi-major axis (1·45 and 1·33 respectively), are much below the significant level of 3. On the other hand, the second harmonic is clearly significant even in the seasonal groups, its ellipses are all far from the origin and the ratios are above 4 for the seasonal means and above 6 for the annual.

On the basis of the above result, the Z-variation is analysed into the first two components only. The results of their significance tests are similar to those of H.

Table 1. Showing the analysis of lunar variation at Ibadan

Seasonal groups of months	C_1 (γ)	t_1 (lunar hour 1st max)	C_2 (γ)	t_2 (lunar hour 1st max)	C_3 (γ)	t_3 (lunar hour 1st max)	C_4 (γ)	t_4 (lunar hour 1st max)
Horizontal force								
D	0·907	9·76	4·696	7·02	0·522	5·70	0·334	0·11
E	1·248	4·73	3·818	7·30	1·322	5·71	0·074	2·24
J	0·989	5·85	3·370	7·30	0·434	2·43	0·224	2·59
A	0·881	6·62	3·988	7·19	0·499	5·52	0·074	0·93
Vertical force								
D	1·093	6·89	3·041	7·21				
E	1·055	11·00	2·072	7·63				
J	0·488	3·10	2·688	7·83				
A	0·612	7·79	2·596	7·52				

2. The Magnitude of L in H at Ibadan

The magnitude of the second harmonic of L in H (3·99 γ) obtained from the analysis of L month by month (Table 1) compares well with the value (3·81 γ) obtained from the analysis at fixed ages of the mean moon (Table 3) despite the fact that not all the data for the first method are included in the second.

As at Huancayo, the second component of L at Ibadan is larger at the December solstice than at the June solstice or the equinoxes. This is expected for Huancayo since it has its summer in December. The difference of 1·3 γ between the seasons is comparable with the maximum error of determination (1·2 γ) but it is still disturbing since both S and L at Ibadan are expected to be larger at the June (summer) than at the December solstice.

The lunar variation in H at Ibadan is very high. Though smaller than the variation at Huancayo, it is about three or more times L in other places of

VOL. 16 1959

comparable latitude. Since L may be expected to be large where S is large, the value of the ratio of S to L is the best test of a real anomaly in L. The ratios are shown in Table 2 (S and L for Bombay are not contemporaneous but the ratio S/L should be sufficiently accurate for the present purpose). The ratios show that L in H at Ibadan as in Huancayo is anomalously large. S/L at Ibadan and Huancayo are about equal and about a third of the value at Bombay.

This is perhaps the first analysis to reveal lunar variation in H with a magnitude comparable to that at Huancayo. It proves that the anomalously

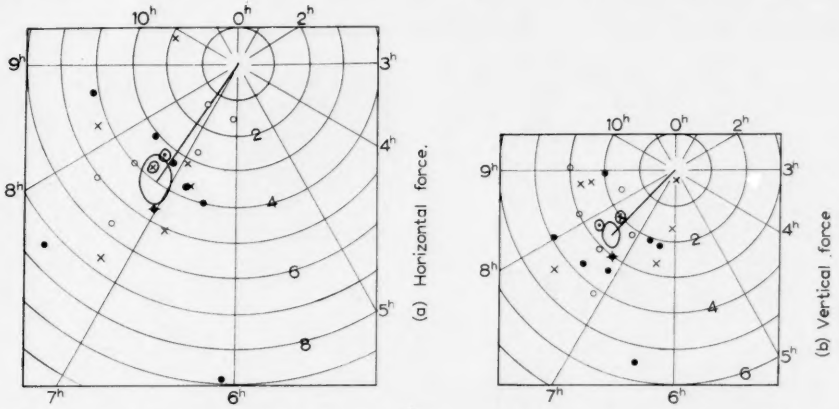


Fig. 1. Showing harmonic dials for the second component of the lunar daily variation of H and Z on undisturbed days at Ibadan; ● December solstice; × equinoxes; ○ June solstice; ⊕ mean December solstice; ⊗ mean equinoxes; ⊙ mean June solstice; radii of circles are in γ.

large L in H at Huancayo is not peculiar to that observatory but probably results from its position near the electrojet and suggests that such large magnitudes of L may be found elsewhere under similar conditions.

There is no significant seasonal variation in phase (Fig. 1). The first maximum occurs at about 7·20 lunar hours after the lower transit of the moon over the longitude of the observatory (1 lunar hour is 1/24 of a lunar day).

3. The Magnitude of L in Z at Ibadan

The second Fourier component of L in Z (2·60 γ) at Ibadan determined from the mean of monthly analysis (Table 1) agrees well with 2·76 γ found from the mean of determinations at eight phases of the moon (Table 3) but they are not derived from exactly the same data.

As in the case of H at Ibadan it is surprising that L in Z is greater at the December than at the June solstice, but again the difference is of the same order as the error of determination. Although there is a progressive retardation in phase from the December to the June solstice (Table 1) the difference of 0·6 hr between them is not as large as the corresponding phase difference in the case of the solar variation. The mean time of the first maximum is 7·52 lunar hours after the lower transit of the moon over the meridian of the observatory.

Table 2. Showing comparisons of S and L at some observatories

Observatory		D months	E months	J months	All months
First component of H					
Ibadan	S				40·96
	L				3·42
	S/L				12·0
Huancayo	S	40·8	51·8	39·6	
(Bartels, 1936)	L	7·1	4·3	2·6	
	S/L	5·7	12·0	15·2	11·0
Bombay	S	16·2	19·6	20·0	
(Schmidt, 1926)	L	1·27	0·88	0·64	
(Chambers, 1887)	S/L	12·8	22·3	31·2	22·1
Second component of H					
Ibadan	S	16·90	21·94	18·99	
	L	4·70	3·82	3·37	
	S/L	3·6	5·7	5·6	5·0
Huancayo	S	23·6	28·5	20·2	
(Bartels, 1936)	L	8·0	5·9	2·4	
	S/L	3·0	4·8	8·4	5·4
Bombay	S	7·9	10·7	10·7	
(Schmidt, 1926)	L	1·35	0·63	0·53	
(Chambers, 1887)	S/L	5·9	17·0	20·2	14·3
First component of Z					
Ibadan	S				18·66
	L				2·14
	S/L				8·7
Second component of Z					
Ibadan	S	11·50	14·86	12·82	
	L	3·04	2·07	2·69	
	S/L	3·8	7·2	4·8	5·3

VOL. 16 1959

The amplitude of L in Z at Ibadan is very large, being over three times the values at Manila and Batavia at which observatories it has been suggested (CHAPMAN, 1951) that Z-variation is anomalously large. This is the largest L in Z so far determined.* The ratio S/L for the second component (Table 2) is equal to the corresponding ratio for H at Ibadan and Huancayo. L in Z at Ibadan is therefore anomalously large.

The analysis of L-variation in Z at Ibadan has shown that the anomalously large amplitude of L first found in H at Huancayo, is not restricted to H or to that observatory. This large amplitude of L must be connected with the existence of the equatorial electrojet near both observatories. The large L-variation in Z at Ibadan almost certainly arises from the position of Ibadan near the latitude of the maximum vertical field of the electrojet.

4. The Chapman Expression for L

CHAPMAN (1913) derived, from the dynamo theory, simple relations for the first four Fourier components of the L-field in terms of the solar time t, the lunar time T, and the lunar age in angular measure v.

$$\left.\begin{aligned} L &= \sum_{n=1}^{n=4} C_n \sin\{(2\pi n/p)T + (n-2)v + \alpha_n\} \\ t &= T + v, \qquad T = 1 \text{ to } p \end{aligned}\right\} \tag{1}$$

This expression is chosen as the basis for the discussion of L at Ibadan because any test of the expression is an indirect test of the dynamo theory. The implications of the expression are examined below.

(a) The amplitudes C_n have no general relations between them. They are constant and therefore independent of the lunar age for which they are computed.

These amplitudes for all four components in H and Z at Ibadan are given in Table 3 together with their standard deviation s. In most cases the means are so many times the standard deviation to allow the conclusion that they are constant. Despite the obvious scatter in a few cases, the same impression is gained from the distribution of the points about a line drawn at the level of their mean (Figs. 2 and 3). One definite conclusion however is that in both H and Z at Ibadan, the amplitudes of all four harmonics show no consistent variation with lunar age.

(b) It follows from the above that for any fixed age of the moon, all four components are real as may be expected from equation (1). Earlier attempts to disregard all but the second led to the assumption that the amplitude depends on solar time and hence the name, luni–solar variation. Because the time each component reaches its first maximum progressively changes with lunar age, the lunar daily variation, for different ages reach their maxima at different times. The curves obtained at Ibadan for both H and Z are similar to those of other observatories. At each age the greatest variation occurs during day-time and

* The comparable magnitude of L in Z at Amberley in New Zealand (BULLEN and CUMMACK, 1953) probably arises from a different cause. The lunar variation in Z at Amberley is, surprisingly, larger than the solar variation.

there is little variation at night. It is also found that the mean variation for the whole lunation is a simple sine wave of period half a lunar day.

(c) The phase angle ϕ_n for the nth harmonic in the Chapman expression for L is

$$\phi_n = (n - 2)v + \alpha_n = 2\pi(n - 2)r/8 + \alpha_n \qquad (2)$$

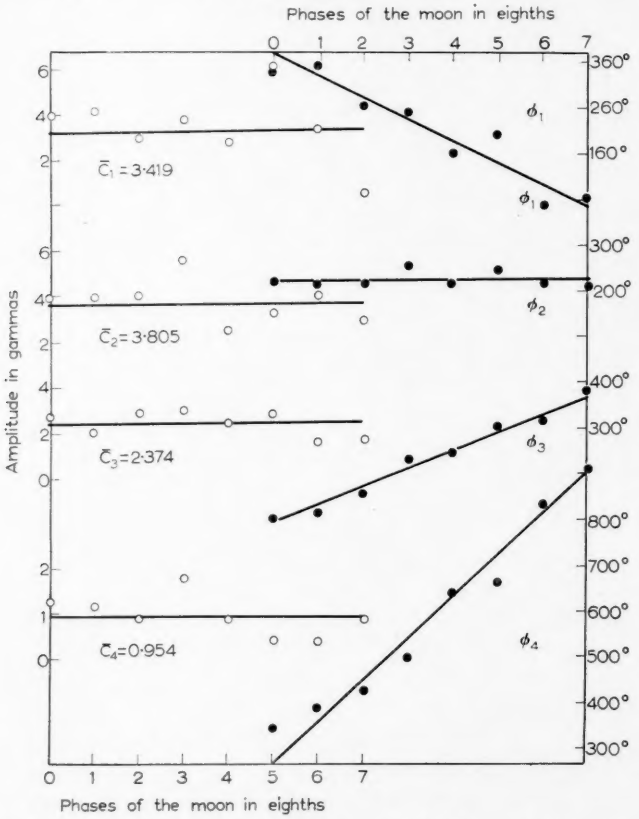


Fig. 2. Showing the amplitudes and phase angles resulting from the harmonic analysis of the lunar daily variation of H at Ibadan, for eight phases of the moon.

where r is the number of the ordinal denoting the eighth (0 for new moon and 7 for seventh-eighth). This expression (2), called the phase law of L, specifies categorically how the phase angle should vary.

In a complete lunation the phase angle of the first harmonic decreases by 2π; the second harmonic has constant phase; the phase angles of the third and fourth harmonics increase by 2π and 4π respectively. The analyses of H and Z at Ibadan are not only in agreement with this but also satisfy the linear relation (2) between ϕ_n and r (Figs. 2 and 3). The theoretical gradients for the four components are -45, 0, $+45$ and $+90$ degrees per eighth of a lunation respectively. The lines of Figs. 2 and 3 give gradients of -50, 0, $+38$ and $+91$ degrees per eighth of a lunation respectively for H and -48, 0, $+48$ and $+94$ degrees respectively for Z. On the whole, there is very good agreement with theory.

VOL. 16 1959

(d) Because of the changes by multiples of 2π all the components but the second should vanish when determined for a lunation.

The analyses of H and Z at Ibadan give ample evidence for this. When the lunar daily variation for eight phases of the moon are averaged, they give a simple sine wave of period half a lunar day. But a clearer evidence comes from the analysis

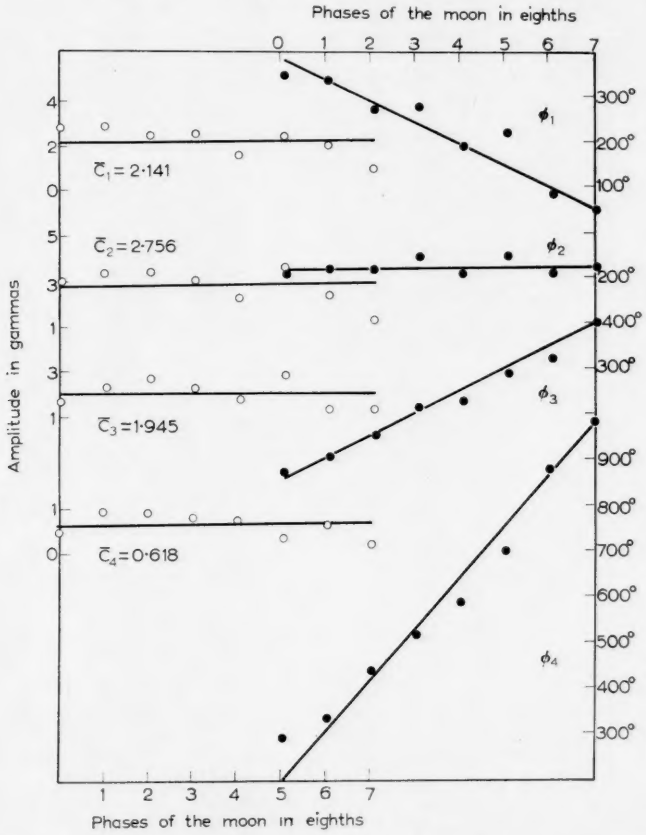


Fig. 3. Showing the amplitudes and phase angles resulting from the harmonic analysis of the lunar daily variation of Z at Ibadan, for eight phases of the moon.

of L, month by month. It has been mentioned that when the four components are statistically tested, only the second is found to be significant. On the other hand, all the four are significant (Table 3) when determined for separate phases of the moon. The same evidence is strengthened in another way. When the determinations at fixed ages are vectorially averaged, the results become of the same order as the errors except in the case of the second harmonic.

(e) The phase constant α_n in the phase law is independent of lunar age. To test this, the experimentally determined ϕ_n is put in equation (2) and using the appropriate values of n and r, α_n is calculated. The standard deviation for the eight values of α_n of each harmonic component is calculated. A table of values

of α_n gives at first an impression of constancy but in fact the standard deviation reveals that in the first and fourth components α_n may vary by 30–40°, and the second and third harmonics by 15–20°. There is however no consistent variation with lunar age.

Table 3. Showing harmonic components of lunar variation in H and Z for eight phases of the moon

Lunar phase r	C_1 (γ)	ϕ_1 (degrees)	C_2 (γ)	ϕ_2 (degrees)	C_3 (γ)	ϕ_3 (degrees	C_4 (γ)	ϕ_4 (degrees)
(*a*) *Horizontal force H*								
0	3·933	349·4	3·970	228·0	2·053	112·5	1·362	350·7
1	4·175	359·4	4·034	219·8	1·987	118·8	1·160	32·7
2	2·887	267·2	4·133	225·6	2·889	163·8	0·903	67·8
3	3·766	257·4	5·594	263·8	3·056	234·3	1·769	138·5
4	2·676	164·9	2·492	217·3	2·536	247·6	0·866	284·8
5	6·084	205·3	3·443	248·1	2·707	303·3	0·412	303·3
6	0·401	52·8	3·936	218·0	1·530	314·1	0·359	118·1
7	3·431	63·8	2·838	207·7	1·626	40·4	0·846	195·5
$\bar{C}$	3·419		3·805		2·375		0·954	
s	1·475		0·457		0·230		0·326	
$\bar{C}/s$	2·30		8·11		10·33		2·93	
(*b*) *Vertical force Z*								
0	2·803	351·7	3·037	214·3	1·554	77·3	0·476	294·6
1	2·806	340·1	3·385	225·0	2·337	107·3	0·950	334·8
2	2·530	275·3	3·418	218·6	2·666	156·8	0·884	80·8
3	2·543	281·3	3·035	252·6	2·216	216·8	0·823	158·6
4	1·536	192·0	2·221	210·2	1·691	226·5	0·672	229·6
5	2·303	220·2	3·513	248·3	2·756	290·0	0·356	341·8
6	1·926	85·2	2·291	210·3	1·195	322·2	0·601	162·5
7	0·781	52·8	1·200	220·4	1·161	42·6	0·182	265·7
$\bar{C}$	2·141		2·756		1·947		0·618	
s	0·680		0·622		0·434		0·230	
$\bar{C}/s$	3·15		4·43		4·48		2·69	

On the whole we conclude that the Chapman (1913) expression for L derived from the dynamo theory can explain the relevant features of L in H and Z at Ibadan.

5. Conclusion

This analysis is important as a confirmation of the abnormally large lunar variation in the horizontal force first observed at Huancayo. It has been shown

VOL. 16 1959

here for the first time that a similar anomally exists in the lunar variation of the vertical force. The results strengthen the belief that the large variations are caused by the effect of the equatorial electrojet which is only about $2\frac{1}{2}^\circ$ N of Ibadan, and probably even nearer Huancayo. The Chapman expression for L has been verified in all aspects. Not many good verifications of the phase law of L exist and this is the first such verification of the law with the vertical force. Any good verification of the phase law inspires confidence in the dynamo theory of L from which it is derived.

Acknowledgements—We are grateful to the International Union of Geodesy and Geophysics for the loan of a set of La Cour variometers and to the Carnegie Institution of Washington for loaning a magnetometer and earth inductor for absolute measurements.

References

Bartels J.	1932	*Terr. Magn. Atmos. Elect.* **37,** 291.
Bartels J.	1936	*Z. Geophys.* **12,** 368.
Bullen J. M. and Cummack C. H.	1953	*J. Geophys. Res.* **58,** 554.
Chambers C.	1887	*Phil. Trans.* A **178,** 1.
Chapman S.	1913	*Phil. Trans.* A **213,** 279.
Chapman S.	1919	*Phil. Trans.* A **218,** 1.
Chapman S.	1951	*Arch. Met., Wien* A **4,** 368.
Chapman S. and Bartels J.	1940	*Geomagnetism* (1st Ed.). Oxford University Press.
Onwumechilli C. A.	1959	*J. Atmosph. Terr. Phys.* This issue, 106.
Schmidt A.	1926	*Arch. Erdmagn.* Bd. I, Heft 1–4. Potsdam.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 124 to 135. Pergamon Press Ltd. Printed in Northern Ireland

The refraction of radio waves by a spherical ionized layer

E. Woyk (E. Chvojková)
Astronomical Institute of the Czechoslovak Academy of Sciences

(*Received* 19 *July* 1958)

Abstract—A general expression is derived for the refraction of radio waves passing completely through a spherical ionized layer, using a simple ray treatment. A similar formula is derived for the ray-path length.

It is shown that when certain approximations appropriate to the terrestrial ionosphere are made, neglecting absorption and the effect of the earth's magnetic field, the expression for refraction reduces to a simple form which has previously been derived by the author using a less general method.

The perpetual propagation of radio waves along two levels in the ionosphere, the condition for which appears in the derivation of the general formula, is briefly discussed.

It is shown that with slight modifications the refraction formula is also valid for reflection, and an expression giving the distance of the point of reflection in the ionosphere from the transmitter on the earth's surface is derived.

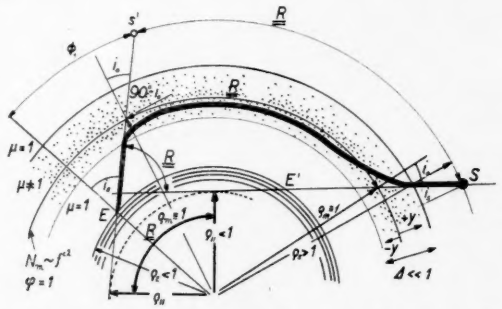


Fig. 1. Propagation of radio waves from cosmic sources. Symbols.

1. Introduction

For radio waves which have been sent from an artificial satellite and which have passed through the whole ionosphere the following equations have been presented (Chvojková, 1958a,b).

$$R = \frac{1}{2}\left(\frac{f^c}{f}\right)^2 \frac{\sin\, i_0}{\cos^3 i_0}\,\sigma \text{ rad} \tag{1}$$

$$d = \overline{ES'} + R\sin i_0 + \gamma\frac{R\cos^2 i_0}{\sin i_0}, \qquad \gamma = 0, \pm 1 \tag{2}$$

These equations are valid for $f^c/f \,.\cos\, i_0 < 0{\cdot}4$ and $\sigma < 0{\cdot}1$. (For $f^c/f.\cos i_0 > 0{\cdot}4$, (14) must be used.)

The meaning of the above symbols, which are shown clearly in Figs. 1 and 2, is as follows:

R is the angle of refraction, or the deviation of a radio wave from its previous direction of propagation, caused by its passing through the ionosphere. (R appears also at the centre of curvature of the penetrated layer.)

VOL. 16 1959

d is the radio path length of the ray from the satellite to the receiver ($\gamma = -1$ applies only when the phase velocity is taken into account, $\gamma = +1$ corresponds similarly to the group velocity).

f^c ($\sim\sqrt{N_m}$) is the critical frequency of the ionosphere (for reflected waves, however, only the critical frequency corresponding to the level of reflection ($\rho_r = 1$) must be used).

f is the frequency of the radio signal.

μ is the refractive index defined by

$$\mu = \sqrt{\{1 - F\varphi(y)\}} \tag{3}$$

where $F = (f^c/f)^2, \qquad \varphi = \varphi(y) = N(y)/N_m$

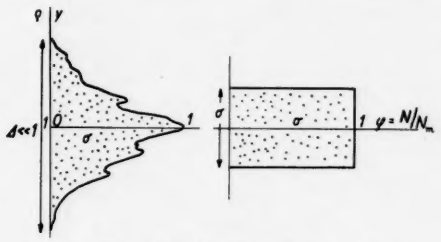


Fig. 2. Equivalent thickness of the layer.

ρ is equal to the radius of curvature of the ionosphere at a given point on the ray path; it is usually the radial distance of that point from the centre of the earth; ρ as well as other lengths are expressed in an appropriate unit. The unit of length chosen is the radial distance of the maximum of electron density N_m (for reflected waves the electron density at the highest level at which the wave can propagate is substituted for N_m and the unit of length is taken to be the radial distance of this level $\rho_r = 1$). Thus

$$\rho \simeq \frac{6400 \text{ km} + h \text{ km}}{6400 \text{ km} + 300 \text{ km}} = 1 + y \tag{4}$$

ρ_m ($= 1$) is the radius of curvature of the ionosphere at N_m.

ρ_E (< 1) is the earth's radius.

ρ_r (< 1) is the radius of the sphere to which the straight part of the ray (outside the ionized layer) must be tangential (Fig. 3). It is given by equation (11), putting $\mu = 1$ and $i = 90°$.

ρ_r ($= 1$) is the level of reflection.

Δ ($\ll 1$) is the thickness of the layer.

y ($= \rho - 1$) is the distance of points under consideration from the level $\rho_m = 1$. Inside the ionosphere $y < \Delta \ll 1$.

i_0 is the angle of incidence at the level $\rho_m = 1$ of the extension of the rectilinear part of the ray path, i.e. it would be the true angle of incidence at this level if there were no ionosphere. i_0 is determined from (11) putting $\mu = 1$ and $\rho = 1$.

σ is called the equivalent thickness of the ionosphere and is defined as the thickness of a layer with the same electron content as the ionosphere but with uniform density N_m (see Fig. 2). In general, for the regions between the levels ρ_1 and ρ_2 we may write

$$\int_{y_1}^{y_2} \varphi(y)\, dy = \frac{1}{N_m} \int_{y_1}^{y_2} N(y)\, dy = \sigma_{12} \tag{5}$$

It should be noted that if $N(y)$ is expressed as electrons per cm^3 (or m^3) then σ is given by

$$\sigma = \int_{-\frac{\Delta}{2}}^{+\frac{\Delta}{2}} N(y)\, dy / N_m \rho_m \tag{5a}$$

where ρ_m is measured in centimetres (or metres) and $\int N(y)\, dy$ represents the total electron content in a column of unit area cross-section.

ES' is the length of a straight line crossing the sphere $\rho_m = 1$ at an angle of incidence i_0 and limited by the levels ρ_E and ρ_S. $ES' = E'S$ would represent the propagation of the ray under consideration if there were no ionosphere between the satellite and the earth's surface.

$$ES' = E'S = \sqrt{(\rho_E{}^2 + \rho_S{}^2 - 2\rho_E \rho_S \cos \phi_1)} \tag{6}$$

ES is the real distance satellite-receiver.

$$ES = \sqrt{\{\rho_E{}^2 + \rho_S{}^2 - 2\rho_E \rho_S \cos (\phi_1 + R)\}} \tag{7}$$

ϕ_1 is the geocentric distance of the point S' from the zenith of E.

$$\phi_1 = \widehat{EOS'} = i_E - i_S \tag{8}$$

$\widehat{EOS}$ is the geocentric distance of the satellite from the zenith of the observer situated at E.

$$\widehat{EOS} = \phi_1 + R \tag{9}$$

$i, i_k, i_E, i_S, i_m, i_0, \ldots$ are the angles of incidence of the ray with the spheres $\rho, \rho_k, \ldots$, i.e. the angles between the ray and the radius vectors at the levels $\rho, \rho_k, \ldots$. From the form of Snell's law for spherical layers, namely

$$\mu\rho \sin i = \text{constant}$$

we have

$$\sin i_k = \frac{\sin i_0}{\mu_k \rho_k}, \qquad \sin i_E = \frac{\sin i_0}{\rho_E}, \qquad \sin i_S = \frac{\sin i_0}{\rho_S} \tag{10}$$

Summarizing, equations (1), (2) and (9) were obtained after having introduced the following arrangement:

(a) The radius of curvature, ρ_m, containing the sphere of maximum electron density of the ionosphere was chosen as the unit of length.

(b) The direction of a ray i_E, i_S, was characterized by the angle i_0 corresponding to the level $\rho_m = 1$. The relation of i_E (at the earth's surface) to i_0 is given by (10).

VOL. 16 1959

(c) The integral (5) was replaced by an equivalent thickness of the layer.
(d) From the appropriate form of Snell's law a certain radius $\rho_{||}$ was found corresponding to the point $i = 90°$ and $\mu = 1$ and given by

$$\mu\rho \sin i = \text{constant} = \sin i_0 = \rho_{||} \tag{11}$$

We see from (11) that those parts of the ray outside the ionospheric layers must always be directed to be tangential to the surface of the same sphere of radius $\rho_{||}$. As a tangent is perpendicular to the corresponding radius-vector, the angle R also appears at the centre of curvature and at the radii $\rho_{||}$, ρ_m, ρ_S, ... (see Figs. 1 and 3).

Thus from the equation (1) (within its limits of validity) and (11) we may make

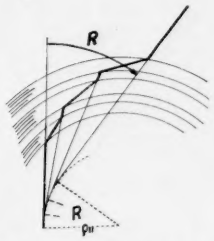


Fig. 3. A spherical layer, divided into a series of spherical elements, separated from each other by non-ionized regions. In these regions the ray must always be tangential to the same sphere $\rho_{||}$ to which it is tangential outside the layer.

the following statement. *Every ray completely crossing a spherical ionized layer will be turned round the centre of curvature of the layer by an angle,* R, *which depends only on* f/f^c, i_0 *and the total electron content of the layer.*

2. Derivation of (1) and (2) and the Most General Formulae

The formula (1) was first derived from

$$\begin{aligned} R &= -\int_{h_1}^{h_2} \frac{d\mu}{\mu} \operatorname{tg} i = -\mu_r\rho_r \int_{\rho_1}^{\rho_2} \frac{d\mu}{d\rho} \frac{d\rho}{\mu\sqrt{(\mu^2\rho^2 - \mu_r^2\rho_r^2)}} \\ &= -\sin i_0 \int_{y_1}^{y_2} \frac{\operatorname{grad} \mu}{\mu} \frac{dy}{\sqrt{\{\mu^2(1+y)^2 - \sin^2 i_0\}}} \end{aligned} \tag{12}$$

developed into Taylor series, assuming that the layer is composed of infinitesimally thin layers inside which the gradient of electron density:

$$\frac{1}{N_m} \operatorname{grad} N = \frac{1}{N_m} \frac{dN}{dy} = \frac{d\varphi}{dy} = \varphi' \tag{13}$$

can be considered constant. An integration over all these thin elements gave (1) and the mentioned limits of its validity. (Chvojková, 1958c–e). After having

reviewed the manuscript Bremmer obtained by a more straightforward derivation the same result but his formula

$$R = \frac{\sin i_0}{2\cos^3 i_0}\left(\frac{f^c}{f}\right)^2 \sum_{m=0}^{\infty}(-1)^m \frac{C_m^{-3/2}}{m+1}\left(\frac{f^c}{f\,.\cos i_0}\right)^{2m}\int \varphi^{m+1}(r)\,dr \tag{14}$$

has enlarged limits, allowing us to calculate the refraction nearly up to the point where the wave must be first reflected (Bremmer, 1958). Next will be used Bremmer's method of derivation for deriving a most general formula, valid even for reflected waves.

Let us introduce $x = \sqrt{\left(1 - \frac{F\varphi}{\cos^2 i_0}\right)}$, (3) and its first derivation

$$\operatorname{grad}\mu = \frac{d\mu}{dy} = -\frac{F\varphi'}{2\mu} \tag{15}$$

into (12) and we have:

$$R = \frac{F\sin i_0}{2}\int_{y_1}^{y_2}\frac{\varphi'\,dy}{\mu^2\sqrt{\{\mu^2(1+y)^2 - \sin^2 i_0\}}}$$

$$= \frac{F\sin i_0}{2}\int_{y_1}^{y_2}\frac{\varphi'\,dy}{(1-F\varphi)[1-F\varphi+(1-F\varphi)(2y+y^2) - \sin^2 i_0]^{1/2}}$$

$$\simeq \frac{F\sin i_0}{2}\int_{y_2}^{y_1}(1-F\varphi)^{-1}\,[\cos^2 i_0 - F\varphi + (1-F\varphi)2y]^{-1/2}\,d\varphi$$

$$= \frac{F\sin i_0}{2\cos i_0}\int_{y_1}^{y_2}(1-F\varphi)^{-1}\left[\left(1 - \frac{F\varphi}{\cos^2 i_0} + \frac{2y}{\cos^2 i_0}(1-F\varphi)\right]^{-1/2} d\varphi$$

$$= \frac{F}{2}\operatorname{tg} i_0\int_{y_1}^{y_2}(1-F\varphi)^{-1}\left(1 - \frac{F\varphi}{\cos^2 i_0}\right)^{-1/2}\sum_{n=0}^{\infty}\binom{-\frac{1}{2}}{n}\left(\frac{2y}{\cos^2 i_0}\frac{1-F\varphi}{1-\frac{F\varphi}{\cos^2 i_0}}\right)^n d\varphi$$

$$= \frac{F}{2}\operatorname{tg} i_0\int_{y_1}^{y_2}\sum_{n=0}^{\infty}\binom{-\frac{1}{2}}{n}\left(\frac{2}{\cos^2 i_0}\right)^n \frac{(1-F\varphi)^{n-1}}{\left(1-\frac{F\varphi}{\cos^2 i_0}\right)^{n+1/2}}\varphi' y^n\,dy = T_0 + \sum_{n=1}^{\infty} T_n$$

$$= T_0 - \sum_{n=1}^{\infty}\binom{-\frac{1}{2}}{n}\frac{\sin i_0}{\cos^{2n-1} i_0}\,2^n\int_{y_1}^{y_2}\frac{(\sin^2 i_0 + x^2\cos^2 i_0)^{n-1}}{x^{2n}}\,x'y^n\,dy$$

$$= T_0 - \sum_{n=1}^{\infty}\sum_{k=0}^{n-1}\binom{-\frac{1}{2}}{n}\binom{n-1}{k}\,2^n\,\operatorname{tg}\,i_0^{2n-2k-1}\int_{t_1}^{\prime} x^{-2n+2k}x'y^n\,dy$$

$$= T_0 + \sum_{n=1}^{\infty}\sum_{k=0}^{n-1}\binom{-\frac{1}{2}}{n}\binom{n-1}{k}\frac{2^n\,\operatorname{tg}\,i_0^{2n-2k-1}}{2n-2k-1}\Big[x^{-(2n-2k-1)}\,y^n - \\ - n\int x^{-(2n-2k-1)}\,y^{n-1}\,dy\Big]_{y_1}^{y_2}$$

VOL. 16 1959

where

$$T_0 = \frac{F}{2}\operatorname{tg} i_0 \int_{y_1}^{y_2} \left(1 - \frac{F\varphi}{\cos^2 i_0}\right)^{-1/2} (1 - F\varphi)^{-1}\,\varphi'\,dy$$

$$= -\left[\operatorname{arctg}\left\{\operatorname{cotg} i_0 \sqrt{\left(1 - \frac{F\varphi}{\cos^2 i_0}\right)}\right\}\right]_{y_1}^{y_2}$$

$$\lim_{\frac{F\varphi}{\cos^2 i_0} < 0{\cdot}1} T_0 = -\left[\sum_{h=0}^{\infty} (-1)^h \operatorname{cotg} i_0^{\,2h+1} \frac{1 - \dfrac{2h+1}{2\cos^2 i_0} F\varphi + \ldots}{2h+1}\right]_{y_1}^{y_2}$$

$$= \frac{F \operatorname{cotg} i_0}{2\cos^2 i_0} \sum_{h=0}^{\infty} (-\operatorname{cotg}^2 i_0)^h [\varphi]_{y_1}^{y_2} = \operatorname{tg} i_0 \left(\frac{f^c}{f}\right)^2 \frac{\varphi_2 - \varphi_1}{2} \qquad (16.1)$$

$$T_1 = -\operatorname{tg} i_0 \left[\frac{y}{\sqrt{\left(1 - \dfrac{F\varphi}{\cos^2 i_0}\right)}} - \int \frac{dy}{\sqrt{\left(1 - \dfrac{F\varphi}{\cos^2 i_0}\right)}}\right]_{y_1}^{y_2}$$

$$\lim_{\frac{F\varphi}{\cos^2 i_0} < 0{\cdot}1} T_1 = \frac{1}{2} \frac{\sin i_0}{\cos^3 i_0} \left(\frac{f^c}{f}\right)^2 \left(\sigma_{12} - [y\varphi\,(y)]_{y_1}^{y_2}\right) \qquad (16.2)$$

$$\text{and} \quad \sum_{n=1}^{\infty} T_n = \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} \sum_{l=0}^{\infty} A_{nkl} \left[y^n \varphi^l - n \int y^{n-1} \varphi^l\,dy\right]_{y_1}^{y_2}$$

$$A_{nkl} = \binom{-\frac{1}{2}}{n} \binom{n-1}{k} \binom{-n+k+\frac{1}{2}}{l} \frac{2^n}{(-1)^l} \frac{\operatorname{tg} i_0^{2n-2k-1}}{2n-2k-1} \left(\frac{f^c}{f \cdot \cos i_0}\right)^{2l} \qquad (16.3)$$

Equations (16.1) and (16.3) lead to the most general formula for refraction, valid for any type of ionized layer surrounding planets or stars and for any sort of propagation, reflection not excluded:

$$R = -\left[\operatorname{arctg}\left(\operatorname{cotg} i_0 \sqrt{1 - \frac{F\varphi}{\cos^2 i_0}}\right)\right]_{y_1}^{y_2} +$$

$$+ \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} \sum_{l=0}^{\infty} \binom{-\frac{1}{2}}{n} \binom{n-1}{k} \binom{-2n+2k+1}{l} \frac{2^n \operatorname{tg} i_0^{2n-2k-1}}{(-1)^l (2_n - 2k - 1)} \left(\frac{f^c}{f \cos i_0}\right)^{2l} \times$$

$$\times \left[\varphi^l y^n - n \int \varphi^l y^{n-1}\,dy\right]_{y_1}^{y_2} \qquad (17)$$

Similarly from

$$d_\gamma = \int_{y_1}^{y_2} \frac{\mu^\gamma\,dy}{\cos i} = \int_{y_1}^{y_2} \frac{\mu^{1+\gamma}(1+y)}{\sqrt{\mu^2(1+y)^2 - \sin^2 i_0}}$$

$$= \sum_{n=0}^{\infty} \binom{-\frac{1}{2}}{n} \frac{2^n}{\cos i_0^{2n+1}} \int \left(\frac{1 - F\varphi}{1 - \dfrac{F\varphi}{\cos^2 i_0}}\right)^{n+1/2} (1 - F\varphi)^{\gamma/2}\, y^n\,dy$$

we obtain the ray-path length d the special case of which is (2):

$$d_\gamma = \sum_n^\infty \binom{-\frac{1}{2}}{n} \frac{2n}{\cos i_0^{2n-1}} \int_{y^1}^{y^2} \frac{(1 - F\varphi)^{n+(1+\gamma)/2}}{\left(1 - \frac{F\varphi}{\cos^2 i_0}\right)^{n+1/2}} y^n \, dy \qquad \gamma = 0, \pm 1 \qquad (18)$$

For the terrestrial ionosphere the following simplifications can be introduced:

(1) As the thickness of the layer, Δ, never much exceeds 0·1 it is sufficient to put $n \leq 1$; consequently $k = 0$ and only (16.1) and (16.2) are to be taken into account.

(2) When $f^c/f < 0{\cdot}4 \cos i_0$ we can put $l = 0$. (When a more precise determination of propagation elements is desired outside these limits, (14) can be used; (17) goes over into (14) for $l = m + 1$. It seems, however, to be rather useless to go outside these limits as then the slightest inhomogeneity in the ionosphere brings about a considerable spreading of the passing wave. We approach to the propagation along the level ρ_∞, see Fig. 5.)

In general three types of propagation must be distinguished:

(a) The propagation through the whole layer; (b) a perpetual run of radiowaves round the earth; (c) reflection at a given level.

3. Refraction of Waves Completely Crossing the Ionized Layer

When we are dealing with waves passing through the whole ionized layer, it

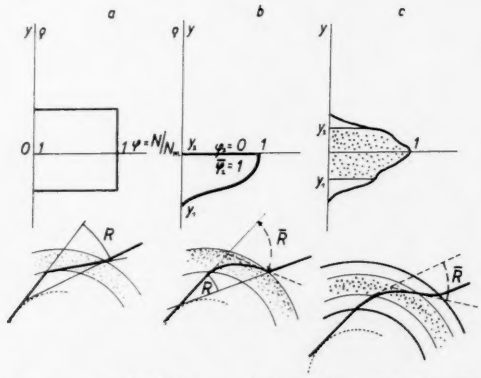


Fig. 4. Some examples showing the difference between the refraction R of a ray completely crossing the layer and the refraction $\bar{R}$ of its part lying completely in the layer.

is sufficient to put $\varphi_1 = \varphi(y_1) = 0$ and $\varphi_2 = 0$ which combined with $n \leqslant 1$ and $l = 0$ leads to (1).

Nevertheless sometimes, although treating waves completely crossing the ionosphere, we are interested to know the deviation of the ray path between two levels only, both lying inside the ionized region. Then, however, at these boundary levels $\varphi_1 \neq 0$, $\varphi_2 \neq 0$, and (16.1) and (16.2) become

$$\bar{R} = \frac{1}{2} \frac{\sin i_0}{\cos^3 i_0} \left(\frac{f^c}{f}\right)^2 [(\varphi_2 - \varphi_1) \cos^2 i_0 + \sigma_{12} - (\varphi_2 y_2 - \varphi_1 y_1)] \qquad (19)$$

VOL. 16 1959

See Fig. 4(b) and (c); it is interesting to consider Fig. 4(a); the refraction of a wave passing through such a homogeneous layer is according to (1):

$$\left[R = \frac{\sin i_0}{\cos^3 i_0}\left(\frac{f^c}{f}\right)^2 \frac{\sigma}{2}\right]$$

while the deviation of the ray in the same limits but inside the layer is according to (19), $\bar{R} = 0$.

4. Perpetual Travel of Radio Waves Round the Earth

In the course of the derivation of (17) the first development into series was prohibited only when

$$\left.\begin{aligned} &\frac{2y(1 - F\varphi)}{\cos^2 i_0 \cdot \left(1 - \dfrac{F\varphi}{\cos^2 i_0}\right)} = -1 \\ &\text{For } y_2 = 0,\ \rho_r \equiv 1,\ \varphi = 1 \\ &\lim_{\substack{y=0 \\ F\varphi = \cos^2 i_0}} \frac{2y(1 - F\varphi)}{\cos^2 i_0 \left(1 - \dfrac{F\varphi}{\cos^2 i_0}\right)} = \frac{-2[(1 - F\varphi) - F\varphi' y]}{F\varphi'} = \frac{2\mu^2}{F\varphi'} = \mu\left(\frac{d\mu}{d\rho}\right)^{-1} \equiv -1 \end{aligned}\right\} \quad (20)$$

The condition (20) is identical with the condition $\mu/\rho = -\,d\mu/d\rho$ derived by Chvojková (1954) in order to define the case when $R = \infty$, i.e. when the wave propagates round the earth.

A detailed analysis there has shown that up to $f \leqslant 8f^c$ there exist in the iono-

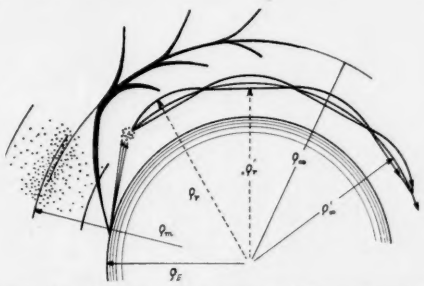


Fig. 5. Two possible modes of propagation round the earth and the difference between them. At $_{\prime\prime}\rho''_r$ the ray is rectilinear.

sphere at least two critical levels at which the gradient of electron density φ' fulfils (20). The upper critical level is the well-known level just below the maximum of electron density of the layer. The lower one appears, however, just over the lower boundary of the layer (usually in a region of negligible electron density where the radius of curvature of the layer equals that of the ray path). The two levels correspond to the minimum and maximum of $\mu\rho$ respectively; see Fig. 6.

As it was shown there, at both these levels the refraction of a horizontally propagating ray becomes infinite.

A wave once tangential to one of these critical levels is obliged to travel along this level round the earth as long as no inhomogeneity of the layer deviates the

wave from the horizontal direction. It was found that the levels possess different character; see Fig. 5. The propagation round the upper critical level is not very stable, as the slightest variation in electron density dissipates the beam in all directions. Once deviated from the horizontal direction the ray never returns back to the upper critical level ρ_∞.

At the lower critical level, ρ_∞', the propagation is very stable. Any inhomogeneity in electron density at this level, deviating the ray from the horizontal direction, forces the wave to come back to this level, so that only oscillations along this level are brought about; see Fig. 5. This perpetual propagation of

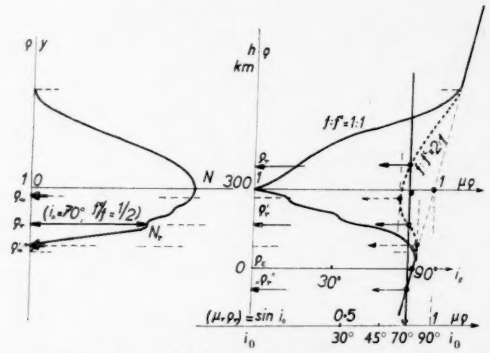


Fig. 6. The $\mu\rho$-diagram (Bremmer, 1949, p. 132; Chvojková, 1954). (*Left*)—The electron distribution with height. (*Right*)—The course of the product $\mu\rho$ plotted to the curve at the left for $f : f^c = 2 : 1$ (dotted curve) and for $f = f^c$ (full curve). One example is pointed out to obtain ρ_r and N_r for $i_0 = 70°$ and $f : f^c = 2 : 1$. The maximum and minimum of $\mu\rho$ correspond to the levels ρ'_∞ and ρ_∞ of Fig. 5, at which a wave of given $f : f^c$ can propagate round the earth.

radio waves along the lower critical level (where no attenuation is produced) would be of great importance for the propagation of radio waves between two very distant stations. Unfortunately either two clouds with increased (or disturbed) electron density must exist at this level (one to bring a part of the transmitted ray into the critical level and the second to bring it back to the second station), or the horizontal variation of electron density must be similar to that treated by Chvojková (1958f). (The levels bounding the mentioned waveguide are easily determined from the $\mu\rho$-curve (see Fig. 6). For $f : f^c = 2 : 1$ and $i_0 = 70°$ a horizontal line through $i_0 = 70°$ crosses the corresponding $\mu\rho$-curve in three points of reflection ρ_r, ρ'_r and $_{''}\rho''_r$. $_{''}\rho''_r$ can be considered as a level of "reflection" in spite of the fact that it is rectilinear. A ray must perpetually reflect between ρ'_r and $_{''}\rho''_r$ as there is a maximum of $\mu\rho$ between them. For $_{''}\rho''_r < \rho_E$ the wave attains the earth's surface. In more complicated cases the layer must be divided into segments as is shown by Chvojková, 1958f).

5. Reflected Waves

The second discrepancy appears in the last part of the derivation of (17), when (16.2) will be developed again into series and when the denominator

$$\sqrt{1 - \frac{F\varphi}{\cos^2 i_0}} = 0$$

VOL. 16 1959

The condition

$$F\varphi = \cos^2 i_0 \tag{21}$$

determines, however, the point where the wave will be reflected. (It is easily found by putting $i_r = 90°$ and $\rho_r = 1$ into equation 11).

A convention must be made before treating the reflected waves: the electron density at the level of reflection ρ_r is the highest electron density which has a meaning for the treated ray. Consequently let us put $\rho_r = {}''\rho''_m \equiv 1$ as well as $\varphi = N/N_r$ and $\varphi_r = 1$. Also i_0 and F must be related to this level.

Thus for reflected waves the previous derivation of (17) remains valid but the value f^c, appearing in $F = (f^c/f)$, must be replaced by the equivalent frequency f_r corresponding to the electron density N_r necessary for reflection.

$$f_r = e\sqrt{\frac{N_r}{\pi m}}$$

The values h_r, ρ_r,N_r of a given wave (defined by i_0) are easily found from the $\mu\rho$-diagram (Fig. 6) discussed by BREMMER (1949) and CHVOJKOVÁ (1954).

As usually $[\rho_r] = 6700$ km, F and i_0 are sufficiently determined by (21), introducing $\varphi = 1$ we have from (11)

$$1 - F = 1 - \left(\frac{f_r}{f}\right)^2 = \sin^2 i_0 = \left(\frac{6400\text{ km} \,.\, \sin i_E}{6400\text{ km} + h_r\text{km}}\right)^2 \simeq \left(\frac{6400}{6700}\right)^2 \sin^2 i_E \tag{22}$$

After this convention the previous derivation of the refraction formula (17) is valid, so that we can proceed up to the formulae (16.1) and (16.2). Thus the refraction of a reflected wave can be expressed by

$$R_r = \frac{\pi}{2} - i_0 + \operatorname{tg} i_0 \left[\frac{-y}{\sqrt{\left(1 - \dfrac{F\varphi}{\cos^2 i_0}\right)}} + \int \frac{dy}{\sqrt{\left(1 - \dfrac{F\varphi}{\cos^2 i_0}\right)}}\right]_{y_1 = -\Delta r \equiv \varphi = 0, F = \cos^2 i_0}^{y_2 = 0 \equiv \varphi = 1, F = \cos^2 i_0} \tag{23}$$

The upper limit is characterized by

$$\left[\sqrt{\left(1 - \frac{F\varphi}{\cos^2 i_0}\right)} = 0,\ y = 0,\ \varphi = 1\right]$$

and the lower one by $y_1 = -\Delta_r$ where Δ_r denotes the distance of the level of reflection from the lower boundary of the layer.

Thus

$$\lim_{\substack{y=0\\ \sqrt{}=0}} \frac{-y}{\sqrt{(1-\varphi)}} = \frac{2\sqrt{(1-\varphi)}}{\varphi'_r} = 0 \qquad \text{and} \qquad -\lim_{\substack{\varphi=0\\ y_1=-\Delta_r}} \frac{-y}{\sqrt{(1-\varphi)}} = -\Delta_r \tag{24}$$

In order to solve the integral of (23) let us first treat a "linear" layer (Fig. 7) having anywhere below ρ_r the same increase of electron density as at the level ρ_r, i.e. $\varphi' = \varphi'_r = \text{constant}$.

Then

$$\int_{y_1}^{y_2} \frac{dy}{\sqrt{(1-\varphi)}} = \frac{1}{\varphi'_r}\int_{\varphi=0}^{\sqrt{}=0} \frac{d\varphi}{\sqrt{(1-\varphi)}} = -\frac{2}{\varphi'_r}\Big[\sqrt{(1-\varphi)}\Big]_{\varphi=0}^{\sqrt{}=0} = 0 + \frac{2}{\varphi'_r}$$

$$= 2\Delta_r^{\text{lin}} \equiv 2\Delta^l_r \tag{25}$$

When the quantity of electrons in the real layer differs from the quantity of electrons in the 'linear'' one, we must consider the surplus or lack of electrons as a layer, completely crossed by the ray. Consequently we suppose them to be squeezed into a secondary equivalent layer the equivalent thickness of which will be denoted by σ_r. A value identical with (1) must be added to (substracted from) the refraction of the "linear" layer. Thus for reflected waves and an arbitrary

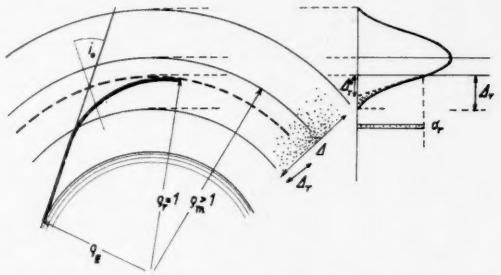


Fig. 7. Reflected waves. The new significance of Δ_r, R_r, ρ_r, σ_r, $\Delta_r{}^l$.

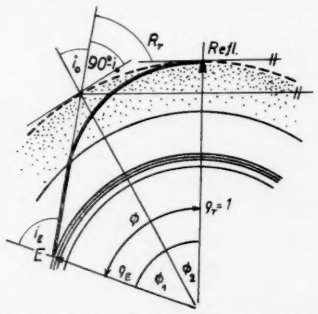


Fig. 8. Reflected waves. Determination of the geocentric distance of the point of reflection ("Refl") from the transmitting station E.

distribution of electrons with height we obtain from (24), (25) and (1) the resulting form

$$R_r = 90° - i_0 + \operatorname{tg} i_0 \left(\Delta^l{}_r \pm \frac{\sigma_r}{2}\right) \tag{26}$$

Equation (26) is not valid only in the close vicinity of the level determined by (20) where $R \to \infty$, and when the gradient of electron density, φ'_r varies substantially even near the reflecting level ρ_r.

The meaning of Δ_r and σ_r follows from Figs. 7 and 8. If expressed in kilometres they must be divided by 6400 km $+ h_r$ km $\simeq$ 6700 km.

The geocentric distance of the point of reflection from the transmitting station is easily determined from Fig. 8.

$$\phi = i_E - i_0 + \frac{180°}{\pi} \operatorname{tg} i_0 \left(\Delta^l{}_r \pm \frac{\sigma_r}{2}\right) = i_E + R_r - 90° \tag{27}$$

VOL. 16 1959

References

Bailey D. K.	1948	*Terr. Magn. Atmos. Elect.* **53,** 41.
Beliayev N. A.	1955	*Astr. J., Moscow* **32,** 359.
Bremmer H.	1949	*Terrestrial Radio Waves.* Amsterdam.
Bremmer H.	1958	*Philips Res. Rep.* In press.
Chvojková E. (Woyk E.)	1954	*Bull. Astr. Insts. Csl.* **5,** 104, 110.
Chvojková E. (Woyk E.)	1958a,b	*Nature, Lond.* **181,** 105, 1195.
Chvojková E. (Woyk E.)	1958c,d,e	*Bull. Astr. Insts. Csl.* **9,** 1, 5, 133.
Chvojková E. (Woyk E.)	1958f	*Nature, Lond.* **182,** 1362.
Hřebík F. and Vokalová E.	1959	*Bull. Astr. Inst. Csl.* **10.** In press.

OL.
16
959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 136 to 144. Pergamon Press Ltd. Printed in Northern Ireland

Polarization computations by means of the multislab approximation*

A. J. FERRARO and J. J. GIBBONS
Ionosphere Research Laboratory, The Pennsylvania State University,
University Park, Pennsylvania

(*Received* 8 *December* 1958)

Abstract—BECKER *et al.* (1951, 1958), discussed a step-discontinuity approximation method for rapidly determining the polarization of low frequency echoes including the effect of collisions and the earth's magnetic field. In the present paper, the results of this method are compared with those of a multislab approximation, and it is shown that the step method yields satisfactory results. For two electron density profiles (one near critical coupling and one far removed) the Rydbeck coupling region was subdivided into slabs of uniform media; by matching the fields between adjacent slabs, $N-1$ simultaneous equations (for N slabs) were obtained. Because of the complexity of these equations, digital computer techniques were employed to obtain the desired solutions. The rate of convergence was estimated by using an increasing number of slab approximations (as high as 10).

1. INTRODUCTION

RECENTLY the state of polarization of long radio wave echoes received from the ionosphere at vertical incidence has been determined by means of a simple formula, BECKER *et al.* (1958), obtained by inserting a discontinuity in the medium; formerly, the "variation of parameters" solution (GIBBONS and NERTNEY, 1952), was used. Both methods are based upon a pair of simultaneous second-order differential equations relating the propagation of a radio wave incident vertically on a horizontally stratified ionosphere (RYDBECK, 1944). These Rydbeck "coupled wave equations" can be written in the form:

$$\begin{aligned}\frac{d^2\pi_0}{dx^2}+[k_0^2N_0^2+M^2]\pi_0 &= -2M\frac{d\pi_x}{dx}-\pi_x\frac{dM}{dx}\\ \frac{d^2\pi_x}{dx^2}+[k_0^2N_x^2+M^2]\pi_x &= -2M\frac{d\pi_0}{dx}-\pi_0\frac{dM}{dx}\end{aligned} \tag{1.1}$$

It is of interest to obtain solutions to (1.1) when M and dM/dx, the so-called coupling factor and its height derivative, respectively, are not negligible and, hence, the two modes are "coupled" to each other.

Briefly, the method of attack for the "variation of parameters" type of solution is as follows:

Given the differential equation

$$y''+P(x)y'+Q(x)y=R(x) \tag{1.2}$$

the two independent solutions (y_{01}, y_{02}) for $R(x)=0$ are first determined. The

* The research reported in this paper has been sponsored by the Geophysics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command, under Contract AF19(604)-3875.

VOL. 16 1959

general solution for any $R(x)$ can be obtained from the variation of parameters in the form:

$$y = Ay_{01} + By_{02} - y_{01}\int_a^x \frac{Ry_{02}}{y_{01}y'_{02} - y'_{01}y_{02}}\,dx + \\ + y_{02}\int_b^x \frac{Ry_{01}}{y_{01}y'_{02} - y'_{01}y_{02}}\,dx \tag{1.3}$$

Equation (1.3) is an exact solution of (1.2); however, the variational method is not exact when applied to (1.1) since the method requires $R(x)$ to be a known function not one containing unknowns as does the right-hand sides of (1.1). The approximate procedure assumes M and dM/dx are equal to zero and the two independent solutions (usually they are W.K.B. functions) are solved for. The zero-order solutions are substituted into the right-hand sides of (1.1) making $R(x)$ a known function; then (1.3) is used to yield first-order solutions (Gibbons and Nertney, 1952).

With the step method, Becker *et al.* (1958), coupling is concentrated by appropriately deforming the electron density vs. height, N–h, and collisional frequency versus height, ν–h, curves in a region around the level of N_c; the deformation is done in such a manner that the gradient of electron density and collision frequency is zero as one approaches the level of N_c from either below or above. At this level there is, then, a step discontinuity in the N and ν profiles. W.K.B. solutions are used as the wave functions and the coupling between the ordinary and extraordinary modes is taken into account by solving a boundary value problem at the discontinuity. This leads to the following formula for the polarization evaluated at the bottom of the step discontinuity:

$$\frac{\pi_x}{\pi_0} = \frac{U_{02} - U_{01}}{1 - U_{01}U_{02}} \tag{1.4}$$

where U_{02} is the Rydbeck polarization term for the ordinary mode evaluated at the top of the discontinuity while U_{01} is evaluated at the bottom. The amplitude and phase of (1.4) must be modified as dictated by the region between the coupling level and the ground level as shown below:

$$\left.\frac{\pi_x}{\pi_0}\right|_{\substack{\text{ground}\\\text{level}}} = \frac{U_{02} - U_{01}}{1 - U_{01}U_{02}} \exp\left\{-k_0\int_0^a (\chi_{x1} - \chi_{01})\,dx\right\} \exp\left\{-jk_0\int_0^a (\mu_{x1} - \mu_{01})\,dx\right\} \tag{1.5}$$

Results from equation (1.5) have been compared with the variational method and they show fair agreement. Since both methods are approximations, it was desirable to have a more accurate method in order to justify the use of (1.5). The multislab solution satisfies this purpose.

2. Multislab Approximation

Russek (1951) discussed the problem for an anisotropic ionosphere in which the slab approximation was used. He considered the general case of oblique incidence upon a horizontally stratified medium and employed a finite number of

slabs in order to obtain a usable approximation. No numerical work was reported in his paper.

This paper employs the multislab approximation for the case of vertical incidence. In the equations which follow, a right-handed co-ordinate system is used with the z-axis magnetic north, y-axis magnetic west and the x-axis positive downwards. The time convention used is $e^{j\omega t}$.

Let us consider consecutive, homogeneous, anisotropic slabs ranging from slab zero to slab N where slab zero extends from $X = 0$ to $X = -\infty$ and slab N from $X = X_{N-1}$ to $X = \infty$ (see Fig. 1). Each slab is characterized by two indices

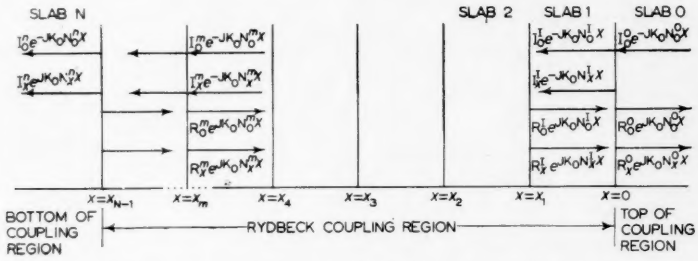


Fig. 1. Multislab approximation for the main echo polarization.

N_x and N_0, the extraordinary and ordinary respectively. Because homogeneous slabs are considered, the coupling factor M and its height derivative dM/dX of (1.1) vanish; therefore, the wave functions are simply upgoing and downgoing plane waves for both the ordinary and extraordinary modes. At the top of the coupling region ($X = 0$) is incident an ordinary wave $I_0{}^0$, which represents the mode returning from the reflection region in that there is no reflection point for the extraordinary mode at low frequencies. In the Nth slab, all waves are downgoing since the Nth slab is outside of the Rydbeck coupling region and no back-scattered signals can be generated. One wishes to solve for the ratio of extraordinary to ordinary in the Nth slab with an ordinary wave exciting the medium at the top of the coupling region.

At every boundary plane separating adjacent slabs, continuity requires tangential $\bar{H}$ and tangential $\bar{E}$ to be continuous; hence, one arrives at four equations by considering the y and z-components of the field quantities and the coupling is taken into account at the boundaries. The electric field components are related to the ordinary and extraordinary modes as follows,

$$\begin{aligned} (E_y)^{\pi_0} &= \{1 - (U_0)^2\}^{-1/2}\pi_0, \qquad (E_z)^{\pi_0} = U_0(E_y)^{\pi_0} \\ (E_z)^{\pi_x} &= \{1 - (U_0)^2\}^{-1/2}\pi_x, \qquad (E_y)^{\pi_x} = U_0(E_z)^{\pi_x} \end{aligned} \tag{2.1}$$

where U_0 is the Rydbeck polarization term for the ordinary mode.
Matching at the Mth interface yields the matrix, after some algebraic manipulation, shown in Fig. 2, where $\overset{*}{I}{}_0^{m+1} = [1 - (U_0^{m+1})^2]^{-1/2}$, etc., and $\underline{\underline{A}}_{nm}$ refers to an element of the array. By means of the total connexion matrix $\underline{\underline{T}}$ for the entire

VOL. 16 1959

$$
\begin{bmatrix} {}^{*}I_0^{m+1} \\ {}^{*}I_x^{m+1} \\ {}^{*}R_0^{m+1} \\ {}^{*}R_x^{m+1} \end{bmatrix} = \begin{bmatrix} N_x^{m+1}(I-U^m U^{m+1})(N_0^m+N_0^{m+1})e^{J(N_0^{m+1}-N_0^m)K_0X_m} & N_x^{m+1}(U^m-U^{m+1})(N_0^{m+1}+N_x^m)e^{J(N_0^{m+1}-N_x^m)K_0X_m} & N_x^{m+1}(I-U^m U^{m+1})(N_0^{m+1}-N_0^m)e^{J(N_0^{m+1}+N_0^m)K_0X_m} & N_x^{m+1}(U^m-U^{m+1})(N_0^{m+1}-N_x^m)e^{J(N_0^{m+1}+N_x^m)K_0X_m} \\ N_0^{m+1}(U^m-U^{m+1})(N_x^{m+1}+N_0^m)e^{J(N_x^{m+1}-N_0^m)K_0X_m} & N_0^{m+1}(I-U^m U^{m+1})(N_x^{m+1}+N_x^m)e^{J(N_x^{m+1}-N_x^m)K_0X_m} & N_0^{m+1}(U^m-U^{m+1})(N_x^{m+1}-N_0^m)e^{J(N_x^{m+1}+N_0^m)K_0X_m} & N_0^{m+1}(I-U^m U^{m+1})(N_x^{m+1}-N_x^m)e^{J(N_x^{m+1}+N_x^m)K_0X_m} \\ a_{13}e^{-J2(N_0^{m+1}+N_0^m)K_0X_m} & a_{14}e^{-J2(N_0^{m+1}+N_x^m)K_0X_m} & a_{11}e^{-J2(N_0^{m+1}-N_0^m)K_0X_m} & a_{12}e^{-J2(N_0^{m+1}-N_x^m)K_0X_m} \\ a_{23}e^{-J2(N_x^{m+1}+N_0^m)K_0X_m} & a_{24}e^{-J2(N_x^{m+1}+N_x^m)K_0X_m} & a_{21}e^{-J2(N_x^{m+1}-N_0^m)K_0X_m} & a_{22}e^{-J2(N_x^{m+1}-N_x^m)K_0X_m} \end{bmatrix} \cdot \frac{1}{2N_0^{m+1}N_x^{m+1}\left[1-(U^{m+1})^2\right]} \cdot \begin{bmatrix} {}^{*}I_0^m \\ {}^{*}I_x^m \\ {}^{*}R_0^m \\ {}^{*}R_x^m \end{bmatrix}
$$

Fig. 2. Matrix relationship for two adjacent slabs.

multilayer, the relationship between waves in slab N and those in slab zero are expressible in terms of the individual matrices $\bar{\bar{T}}^{(m)}$ (as in Fig. 2) by:

$$\bar{C}^{(N)} = \bar{\bar{T}}\,\bar{C}^{(0)} = \prod_{R=0}^{N-1} \bar{\bar{T}}^{(R)} \bar{C}^{(0)} \tag{2.2}$$

For the determination of the reflected echo polarization $\bar{C}^{(N)}$ and $\bar{C}^{(0)}$ have the forms

$$\bar{C}^{(N)} = \begin{bmatrix} \overset{*}{I}_0{}^N \\ \overset{*}{I}_x{}^N \\ 0 \\ 0 \end{bmatrix}, \qquad \bar{C}^{(0)} = \begin{bmatrix} \overset{*}{I}_0{}^0 \\ 0 \\ \overset{*}{R}_0{}^0 \\ \overset{*}{R}_x{}^0 \end{bmatrix} \tag{2.3}$$

Once $\bar{\bar{T}}$ is determined, only four unknowns (and four equations) remain, transmitted ordinary and extraordinary at the bottom of the coupling region and reflected ordinary and extraordinary at the top of the coupling region. Any amplitude can be assumed for $\overset{*}{I}_0{}^0$ since one desires only the ratio of:

$$\frac{\overset{*}{I}_x{}^N}{\overset{*}{I}_0{}^N}$$

the complex polarization ratio.

VOL. 16 1959

3. Digital Computer Application to the Multislab Approximation

No simple analytical solution is possible even for a limited number of slabs, except for two (Becker *et al.*, 1958), because of the complexity of the elements of $\bar{\bar{T}}^{(m)}$, which involve real and imaginary terms, and in addition because of the lengthy expressions resulting from matrix multiplications. Only numerical examples can be treated; even then, the work would be too tedious for hand-computation. For this reason, use was made of "Pennstac", a digital computer in operation at the Pennsylvania State University.

Before beginning the computer solution, one must compute the indices and polarization terms for all the slabs; since the difference of the indices occur as terms in the matrix, one must compute these to a large number of significant figures in that this difference in the coupling region is extremely small. For the computations presented in this paper, N_x, N_0 and U_0 were hand-computed to eight significant figures; however, there is now available at this Laboratory a computer programme to compute these terms rapidly on "Pennstac."

A total of 550 instructions were used in the computer programme to compute the elements, store them and then perform the multiplications; total computing time is $4{\cdot}5\,N$ min where N is the number of interfaces used. The factor post multiplying the individual matrices (see Fig. 2) was not computed since it ultimately cancels itself in the ratio

$$\frac{\overset{*}{I}_x{}^N}{\overset{*}{I}_0{}^N}$$

4. Application of the Multislab Approximation to the Ionosphere

The multislab approximation was applied to two models, one exhibiting strong coupling, the other weaker coupling. The first model studied was that of Parkinson (1955) in which the N_c level and ν_c level are only 0·64 km apart, hence producing strong coupling. Fig. 3 shows this model along with the collision frequency profile (Nicolet, 1953). The top of the coupling region $X = 0$ represents a height of 90·5 km. Solutions were obtained for 2, 3, 5 and 9 slab approximations; these slabs are shown only on the N profile but must be included in the ν profile as well. The results for 150 kc/s are summarized in Table 1 and displayed in Fig. 4.

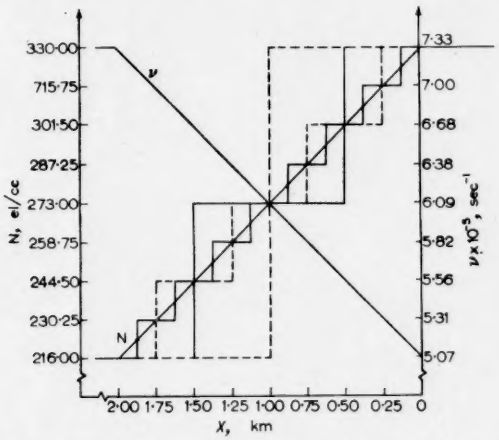


Fig. 3. Models of electron density and collision frequency for almost critical coupling conditions.

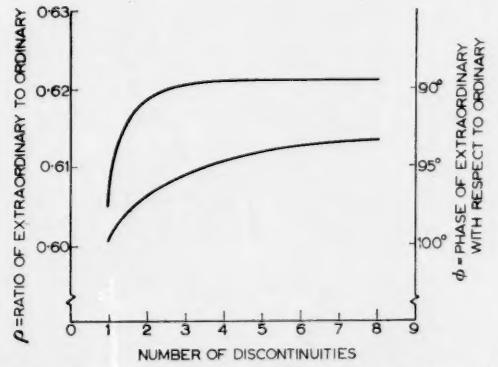


Fig. 4. Polarization of main echo by multislab method for near critical coupling conditions.

Equation (1.4) yields the same result as the one discontinuity of Table 1; it was included in the computer solution only as a check upon computer operation. As can be seen, the rapid method (1.4) compares favourably with the multislab approximation.

Table 1. Polarization of the reflected echo for near critical coupling conditions

Number of discontinuities	Magnitude of polarization ratio	Phase of polarization ratio
1	0·6050	−99·77°
2	0·6194	−96·80°
4	0·6220	−94·70°
8	0·6221	−93·60°

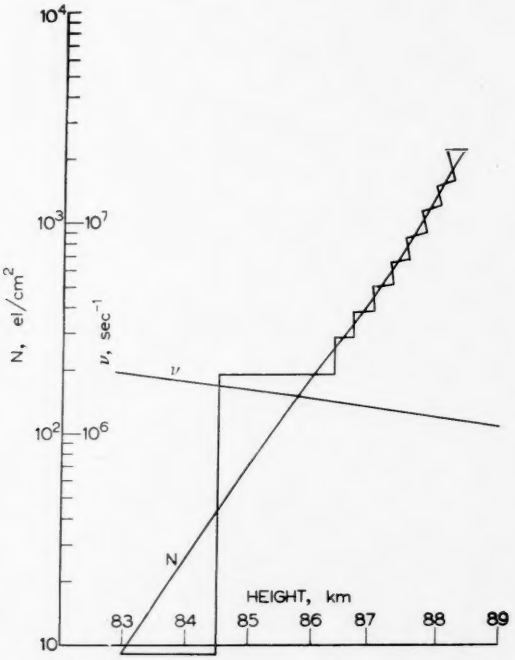


Fig. 5. Chapman model and collision curve for far from critical coupling conditions.

For weaker coupling conditions, a Chapman model with a penetration frequency of 4·4 Mc/s, (Nertney, 1951), was used. The N and ν profiles are expressed by:

$$N(\text{el/cm}^3) = 2{\cdot}3935 \times 10^5 \,.\, \exp \tfrac{1}{2}(1 + x - e^x)$$

$$\nu(\text{sec}^{-1}) = 0{\cdot}8 \times 10^5 \,.\, e^x \qquad (4.1)$$

where

$$x = \frac{115 - h\,(\text{km})}{10}$$

and are shown in Fig. 5. Although only 10 slabs are shown in the figure, solutions were also obtained for 3, 4 and 6 slabs. Only 2 slabs are necessary between 83 km and 86·5 km since the electric polarization term is so slowly varying that two slabs

VOL. 16 1959

produce a good approximation. Results for 150 kc/s are tabulated in Table 2 and plotted in Fig. 6.

NERTNEY (1951) obtained $0{\cdot}3 \angle +90°$ by the variational method, the sign difference being due to his choice of co-ordinate axis. Equation (1.4) gives $0{\cdot}17 \angle -61°$.

Greatest discrepancy between the multislab and rapid approximation is found

Table 2. Polarization of the reflected echo for far from critical coupling conditions

Number of discontinuities	Magnitude of polarization ratio	Phase of polarization ratio
2	0·1821	−80·80°
3	0·2559	−87·90°
5	0·2701	−90·06°
9	0·2863	−91·42°

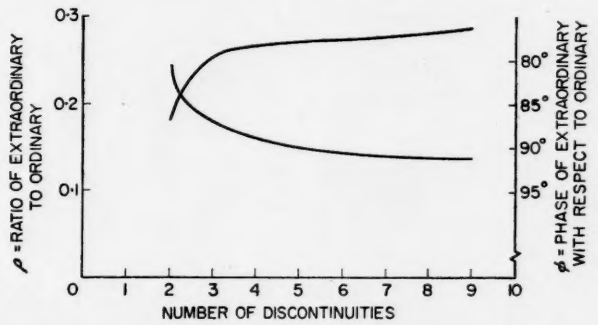


Fig. 6. Polarization of main echo by multislab method for far from critical coupling conditions.

for this Chapman model which produces weak coupling; this is easily explained. With strong coupling conditions (N_c level near this ν_c level) the polarization term U changes rapidly in a region around the N_c level. Its variation with height closely resembles a step jump; therefore, a single discontinuity is a fair approximation. With weaker coupling conditions, its variation is more gradual so that several slabs are required for a good fit. Still, equation (1.4) gives reasonable results.

5. Conclusion

The aim of the paper has been to justify the use of a fast, simple technique for dealing with the polarization problem; such a scheme would be readily applicable to determining the long-wave, sweep-frequency polarization for a given ionosphere model.

For quick analysis of models, one can apply the rapid method, or for more detailed study rely upon the computer technique of the multislab problem.

6. References

Becker W.	1951	*Z. Angew. Phys.* **3,** 83.
Becker W., Ferraro A. J. and Gibbons J. J.	1958	Scientific Report No. 104. Ionosphere Research Laboratory, The Pennsylvania State University, University Park, Pennsylvania.
Gibbons J. J. and Nertney R. J.	1952	*J. Geophys. Res.* **3,** 323.
Nicolet M.	1953	*J. Atmosph. Terr. Phys.* **3,** 200.
Parkinson R. W.	1955	*J. Atmosph. Terr. Phys.* **7,** 173.
Russek A.	1951	Research Report No. EM-38. Mathematics Research Group, New York University.
Rydbeck O. E.	1944	Report No. 7. Research Laboratory of Electronics, Chalmers University of Technology.

List of Symbols

$k_0^2 = \dfrac{\omega^2}{C^2}$, where C is in km/sec.

π_0, π_x = the two "Rydbeck" modes—ordinary and extraordinary.

N_0, N_x = the refractive indices associated with the above modes.

U_0, U_x = electric polarization terms for the above modes.

$M = \dfrac{dU}{dx} / (1 - U^2)$ = coupling factor.

μ = real part of the refractive index.

χ = imaginary part of the refractive index.

N_c = 278·454 el/cm³ for 150 kc/s, i.e. $\dfrac{4\pi N_c e^2}{m\omega^2} = 1$

where e = charge of electron

m = mass of electron

ω = operating angular frequency.

ν_c = 5·348 × 10⁵/sec at State College, i.e. $\dfrac{\omega_H^2}{\sigma^2 \nu_C^2} = 1$

where ω_H = gyro-angular frequency

$\sigma = 2 \cos \theta_p / \sin^2 \theta_p$

θ_p = angle between direction of propagation and earth's magnetic field.

I_0^m, I_x^m = downgoing ordinary and extraordinary modes in slab m.

R_0^m, R_x^m = upgoing ordinary and extraordinary modes in slab m.

$\bar{\bar{T}}$ = total connexion matrix between waves in slab N and slab 0.

$\bar{\bar{T}}^{(m)}$ = matrix relationship between two adjacent slabs.

$\bar{C}^{(N)}$, $\bar{C}^{(0)}$ = column vectors representing waves in slabs N and 0 respectively.

VOL. 16 1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 145 to 146. Pergamon Press Ltd. Printed in Northern Ireland

Rhythmic fading of short-wave radio signals

B. N. Singh and R. L. Ram
Physics Laboratory, Patna University, Patna

(*Received* 22 *November* 1958)

Abstract—The problem of periodic fading due to interference of a number of components of a sinusoidal radio signal has been theoretically treated by considering it as equivalent to determining the resultant of a number of simple harmonic vibrations of nearly equal frequencies. The practical bearing of the results so obtained to fading curves of magneto-ionic origin particularly those caused by interference of three and four components has been discussed. A number of typical fading curves for short-wave radio signals transmitted from Delhi and Calcutta have been described. Particular reference may be made of curves which are very likely of magneto-ionic origin and depict the complete sequence of events expected when the value of the m.u.f. for the *F2*-layer passes across the signal frequency due to the increase or decrease of the ionic density of the layer. Fading curves which have been ascribed to interference between waves singly and doubly reflected from *F2*-layer and also those which are apparently caused by interference between waves simultaneously reflected from *F2*- and *E*-layers have also been described and their features discussed. An example has also been given of such fading curves whose origin appear to be obscure.

1. Introduction

The rhythmic or periodic fading of radio signals is supposed to be produced by interference between two or more components which may belong to any one of the following categories: (a) components of magneto-ionic origin; (b) components reflected singly and multiply from the same ionospheric layer; (c) components reflected simultaneously from two layers; (d) direct (ground) ray and any one of the components reflected from the ionosphere. It is quite possible that some periodic fading may be produced due to interference between components which belong to any two of the categories mentioned above.

The periodic fading curves formed by interference between magneto-ionic components are usually observed during morning or evening hours when the ionic density of the ionospheric layers is changing and they are likely to become more conspicuous when the frequency of the radio signal approaches m.u.f. for the given layer and for the receiving station concerned. Appleton and Beynon (1947) have described the sequence in which the rhythmic pattern would appear in such cases as far as reflection from *F2*-layer is concerned. They have also reported the occurrence of the complete sequence of the above type of fading during evening hours. Fading curves of magneto-ionic origin have also been observed by other workers (Banerjee and Singh, 1948, 1949; Khastgir and Das, 1950; Khastgir and Tantry, 1951) though no one has reported the full sequence of events expected near the m.u.f. for the layer concerned. Recently Singh and Ram (1958) have reported a peculiar type of fading which they attribute to interference of as many as four magneto-ionic components.

Periodic fading due to interference between waves either singly and doubly reflected from the same ionospheric layer or simultaneously reflected from two ionospheric layers has been studied by several workers (Banerjee and Mukherji,

1946, 1948; Banerjee and Singh, 1948, 1949; Khastgir and Das, 1950; Khastgir and Tantry, 1951). Many fading patterns have been attributed to interference of magneto-ionic components reflected from one layer and a component reflected from another layer (Khastgir and Das, 1950; Khastgir and Tantry, 1951; Singh and Ram, 1958).

In the present paper the problem of periodic fading produced by interference of two or more components of a sinusoidal radio signal has been considered as equivalent to finding out the resultant of the e.m.f.'s of two or more radio-frequency simple harmonic vibrations of nearly equal frequencies. From this view point the nature of the periodicity resulting from interference between two, three and four components have been discussed. Attention has been drawn to the fact that periodic fading of magneto-ionic origin may also arise on account of a vertical motion of a reflecting region. The paper also contains some typical periodic fading curves recorded by the authors for short-wave radio signals transmitted from Delhi and Calcutta. The fading curves of magneto-ionic origin illustrate the full sequence of expected events not only during the fall in the electronic density of the *F2*-layer but also when the same is rising. Fading patterns of other origins have also been described. The nature of some of the fading curves has been discussed in the light of the conclusions arrived at in the theoretical part of the paper.

2. Theoretical Considerations

When two components of the same radio signal traversing different paths interfere, they would produce a periodic fading if they are of comparable amplitudes and their equivalent path difference varies at a constant rate with time. If the equivalent path traversed by either of them be changing with time at a constant rate it would produce an apparent increase or decrease in the frequency of the component depending upon whether the equivalent path is decreasing or increasing. Thus the problem of interference between two components of a radio signal cited above may be considered as equivalent to the consideration of the resultant of the e.m.f.'s of two radio waves of slightly different frequencies, the difference being directly proportional to the rate of variation of the equivalent path difference of the two components. If the apparent frequencies of the components be represented by $\omega_1/2\pi$ and $(\omega + \omega_1)/2\pi$ (the apparent frequency of one of them being assumed to remain the same as the frequency of the signal without any loss of generalization) and the e.m.f.'s by $a \sin \omega t$ and $a_1 \sin (\omega + \omega_1)t$ respectively, the resultant will be given by:

$$E = R \sin (\omega t + \alpha) \tag{1}$$

where

$$R^2 = a^2 + a_1^2 + 2aa_1 \cos \omega_1 t \tag{2}$$

and

$$\tan \alpha = \frac{a_1 \sin \omega_1 t}{a + a_1 \cos \omega_1 t} \tag{3}$$

As $\omega \gg \omega_1$ the term $\sin (\omega t + \alpha)$ will remain a radiofrequency term and will, therefore, not appear as a periodic variation in the output of a radio-receiver, but the variation in the amplitude of frequency $\omega_1/2\pi$ will be present in the output.

VOL. 16 1959

If interference takes place between three radio waves and the path difference between any two of them be changing continuously at a constant rate with time, the resultant e.m.f. may be represented by:

$$E = a \sin \omega t + a_1 \sin (\omega + \omega_1)t + a_2 \sin (\omega + \omega_2)t \tag{4}$$

where a, a_1 and a_2 are the amplitudes and $\omega/2\pi$, $(\omega + \omega_1)/2\pi$ and $(\omega + \omega_2)/2\pi$ the apparent frequencies of these waves respectively.

The equation (4) may be written as

$$E_1 = R_1 \sin (\omega t + \beta) \tag{5}$$

where

$$R_1^2 = a^2 + a_1^2 + a_2^2 + 2aa_1 \cos \omega_1 t + 2aa_2 \cos \omega_2 t + 2a_1 a_2 \cos (\omega_2 - \omega_1)t \tag{6}$$

and

$$\tan \beta = \frac{a \sin \omega_1 t + a_2 \sin \omega_2 t}{a + a \cos \omega_1 t + a_2 \cos \omega_2 t} \tag{7}$$

As $\omega \gg \omega_1$ or ω_2, $(\omega t + \beta)$ will be a high-frequency term and it will not produce a periodic variation in the output. The nature of the variation of the amplitude will be given by equation (6) from which it is evident that there are as many as three low-frequency periodicities of magnitudes $\omega_1/2\pi$, $\omega_2/2\pi$ and $(\omega_2 - \omega_1)/2\pi$ present in the output. If a_1 and a_2 are small compared to a, only two periodicities of frequencies $\omega_1/2\pi$ and $\omega_2/2\pi$ will be effective in varying the amplitude of the signal. If a_1 and ω_2 are sufficiently large than a_2 and ω_1 respectively, fast rhythmic variations of a small amplitude may appear superimposed upon a slow one of large amplitude.

If ω_1 is very nearly equal to ω_2, it would be profitable to put the equation (6) in the form

$$R_1^2 = a^2 + a_1^2 + a_2^2 + 2a \sqrt{\{a_1^2 + a_2^2 + 2a_1 a_2 \cos (\omega_2 - \omega_1)t\}} \cos (\omega_1 t + \gamma) + \\ + 2a_1 a_2 \cos (\omega_2 - \omega_1)t \tag{8}$$

where,

$$\tan \gamma = \frac{a_2 \sin (\omega_2 - \omega_1)t}{a_1 + a_2 \cos (\omega_2 - \omega_1)t}$$

If $a \gg a_1$ or a_2, the above expression reduces to

$$R_1^2 = a^2 + 2a \sqrt{\{a_1^2 + a_2^2 + 2a_1 a_2 \cos (\omega_2 - \omega_1)t\}} \cos (\omega_1 t + \gamma) \tag{9}$$

From this equation it is clear that in this case the amplitude would vary in such a way as to present the appearance of a beat of frequency $\omega_2 - \omega_1$. Even if a is not much greater than a_1 or a_2, still the beat form may appear superimposed upon another rhythmic variation in the output of the receiver having the same frequency as that of the beat.

Let us now consider four r.f. waves whose e.m.f.'s are represented by $a \sin \omega t$, $a_1 \sin (\omega + \omega_1)t$, $a_2 \sin (\omega + \omega_2)t$ and $a_3 \sin (\omega + \omega_3)t$.
Their resultant will be represented by

$$E_2 = R_2 \sin (\omega t + e) \tag{10}$$

where, R_2 and e are fairly lengthy expressions involving a, a_1, a_2, and a_3 besides ω, ω_1, ω_2 and ω_3; $\sin(\omega t + e)$ will, however, remain a h.f. term and hence as usual will not vary the output of the receiver. The latter will vary according to R_2. If a is sufficiently large compared to a_1, a_2 or a_3 we shall get approximately

$$R_2^2 = a^2 + 2aa_1 \cos \omega_1 t + 2aa_2 \cos \omega_2 t + 2aa_3 \cos \omega_3 t \tag{11}$$

The nature of this expression will be a complicated one but in special cases it may assume simpler forms. If ω_2 and ω_3 are nearly equal we can put the equation (11) in the following form:

$$R_2^2 = a^2 + 2aa_1 \cos \omega_1 t + 2aQ \cos\left(\frac{\omega_2 + \omega_3}{2} + \delta\right) \tag{12}$$

where,

$$Q^2 = a_2^2 + a_3^2 + 2a_2 a_3 \cos(\omega_2 - \omega_3)t \tag{13}$$

and

$$\tan \delta = \frac{a_2 - a_3}{a_2 + a_3} \tan(\omega_2 - \omega_3)t \tag{14}$$

If ω_2 or ω_3 is large compared to $\omega_2 - \omega_3$ the expression $2aQ \cos\{(\omega_2 + \omega_3)/2 + \delta\}$ will present the appearance of a beat of frequency $\omega_2 - \omega_3$. If ω_1 is further sufficiently smaller than ω_2 or ω_3, and a_1 is sufficiently greater than a_2 or a_3, the beat pattern described above would appear superimposed upon the slow periodicity of a larger amplitude.

We may consider the application of the results derived above to cases of interference between the magneto-ionic components. When the frequency of a radio signal is lower than the m.u.f. and the upper trajectory of the ordinary wave is absent, we have to consider the resultant of only two components, that is the lower trajectories of the ordinary and extraordinary waves. Similarly, when the frequency of the signal is greater than the m.u.f. for the ordinary ray and less than that for the extraordinary, there will be only two components, viz. the upper and lower trajectories of the extraordinary. In both of these cases the resultant amplitude will be periodic as shown by equation (2) provided that the equivalent path difference between the two interfering components varies at a constant rate. If the rate of variation of the path difference, instead of being constant, changes steadily a periodicity of varying frequency may be obtained. When the frequency of the signal is less than the m.u.f. for the ordinary component and close to it, the upper trajectory ray of the ordinary may also be present. In this case if we represent the e.m.f.'s of the lower trajectory of the ordinary, the lower trajectory of the extraordinary and the upper trajectory of the ordinary by $a \sin \omega t$, $a_1 \sin(\omega + \omega_1)t$ and $a_2 \sin(\omega + \omega_2)t$ respectively as in equation (4), the resulting fading pattern will be a complicated one as represented in equation (6). Usually we may consider a to be much greater than a_1 or a_2 and in most cases a_1 may be considered as rather greater than a_2. In such cases provided we are not very close to the m.u.f. for the ordinary wave we can further assume that ω_2 is sufficiently greater than ω_1. If interference of the three components takes place under conditions mentioned above,

VOL. 16 1959

the resultant fading pattern would consist of a fast rhythmic variation of small amplitude superimposed over a slow periodic curve of a larger amplitude. When the signal frequency is very close to m.u.f. for the ordinary wave at some stage the periodic pattern would take the form of a beat as represented by equation (9). If the upper trajectory ray also appears along with the other three components, the former may be represented by $a_3 \sin(\omega + \omega_3)t$ and then the resultant interference pattern will be according to the equation (11). If now, $a \gg a_1, a_2$, or a_3 and a_1 is also sufficiently greater than a_2 or a_3 (which is likely) and if further ω_2 or $\omega_3 > \omega_1$ and ω_2 or $\omega_3 > \omega_2 - \omega_3$, the resultant amplitude will be of a beat pattern superimposed upon a slow periodic one. The above conditions are likely to be fulfilled if the frequency of the signal is slightly less than the m.u.f. for the ordinary wave but is not too close to it. The fading pattern of this type has been reported by Singh and Ram (1958).

In case of interference between rays singly and doubly reflected from the same ionospheric layer, by applying results of equation (2) we find that the resultant fading pattern would be periodic of frequency $\omega_1/2\pi$. The latter can be further shown to be equal to:

$$n = \frac{2v}{\lambda}(2\cos\theta_2 - \cos\theta_1) \tag{15}$$

where θ_1 and θ_2 are the angles of incidence for the waves singly and doubly reflected from the layer which is moving with a vertical velocity v and λ is the wavelength of the radio signal. It may be mentioned that the equivalent heights of the points of reflection for the ordinary and extraordinary components of a signal are different and hence even when the electronic density of the reflecting region be assumed to remain constant, a periodic fading may be produced by the vertical movement of the reflecting layer. In view of the fact that the equivalent separation between the two points of reflection is usually not large the frequency of the periodic fading would be generally of a small magnitude except when the signal frequency is close to the m.u.f. for the reflecting layer. In the latter case the same sequence of magneto-ionic pattern as produced by the variation of electron density near the m.u.f. will be expected.

If interference takes place between waves simultaneously reflected from two different ionospheric layers moving vertically, the frequency of the periodic fading may be shown to be given by

$$n_1 = \frac{2}{\lambda}(v_1 \cos\phi \pm v_2 \cos\psi) \tag{16}$$

where, ϕ is the angle of incidence of the wave reflected from the higher layer moving with a vertical velocity v_1 and ψ is the angle of incidence of the wave reflected from the lower layer which is moving with a vertical velocity v_2; a plus sign will be used when the layers are moving in opposite directions and a minus sign when they are moving in the same direction.

It would not be out of place to mention that in case of reflections from the *E*-layer the sequence of magneto-ionic fading near the m.u.f. may not be the same as that for the *F2*-layer under similar conditions. For example, let us consider

the sequence when the electronic density of the *E*-layer is falling and the m.u.f. for this layer for a given distance between the transmitter and receiver is approaching the signal frequency. First the periodic pattern due to interference between the lower trajectory rays of the ordinary and extraordinary waves may occur and then other phases of the pattern may follow but before the full sequence of events has taken place simultaneous reflection from the *F2*-layer may also appear on the scene with the result that the periodic pattern may become too complicated. A similar thing may happen at the time of the formation of the *E*-layer.

3. Experimental Arrangement

The experimental arrangement made by the authors was of the usual type. The output of the second detector of a superhet receiver with one r.f. stage of amplification and without any a.v.c was connected to a multiflex moving-coil galvanometer, the reflected light spot of which controlled the movement of the pen of a recorder by a photo-electric device. The fading curves were obtained on a tracing paper. Fading patterns for short-wave transmissions from various stations of All India Radio were recorded for different hours of day and night for several months. The aerial used was of a vertical type and the receiver was tested to see that its output was linear.

4. Experimental Results and Discussion

Curves in Figs. 1 and 2 may be considered as fading patterns of magneto-ionic origin. They are for a radio signal of wavelength 31·3 m transmitted from Delhi which is at a distance of about 1000 km from the receiving station. During the period of observation the electronic density of the *F2*-layer may be supposed to be steadily falling which is usual during the night-time. It would also be not unreasonable to assume that by 0024 hours the m.u.f. for the *F2*-layer for the given transmitting and receiving stations had sufficiently decreased to come fairly close to the frequency of the signal. The slow periodic fading of Fig. 1(a) may then be considered as caused by interference between the lower trajectory rays of the ordinary and extraordinary waves. As the m.u.f. for the ordinary wave comes closer to the frequency of the signal the frequency of the slow periodicity increases as shown in Fig. 1(b). The quick rhythmic pattern superimposed upon the slow periodicity of larger amplitude as shown in the same figure can be ascribed to interference between the upper and lower trajectory rays of the ordinary wave. When the ordinary wave is on the point of penetrating the *F2*-layer its intensity is expected to increase suddenly and this is perhaps illustrated in Fig. 1(c). When the ordinary wave has ceased to be reflected the intensity of the signal should fall and this fact is also evident in the same figure. The periodic fading of Figs. 1(d) and 2(a) can be due to interference of the lower and upper trajectory rays of the extraordinary wave. The disappearance of the extraordinary wave is perhaps indicated in Fig. 2(a). The weak signal of random type left afterwards is very likely due to back scattering. Its feebleness and intermittence confirms this conclusion. The sudden increase in the signal intensity shown in Fig. 2(b) is obviously due to some temporary spurious reflection as it is of a very short duration (Figs. 2c and 2d).

VOL. 16 1959

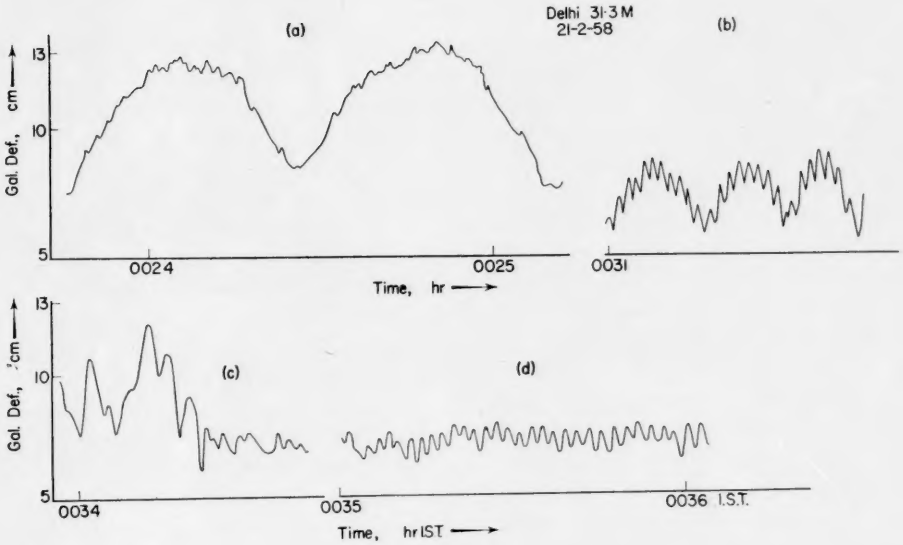

Delhi 31·3 M
21-2-58
(a)
(b)
(c)
(d)
Gal. Def., cm
13
10
5
0024
0025
0031
Time, hr
0034
0035
0036 I.S.T.
Time, hr I.S.T

Fig. 1.

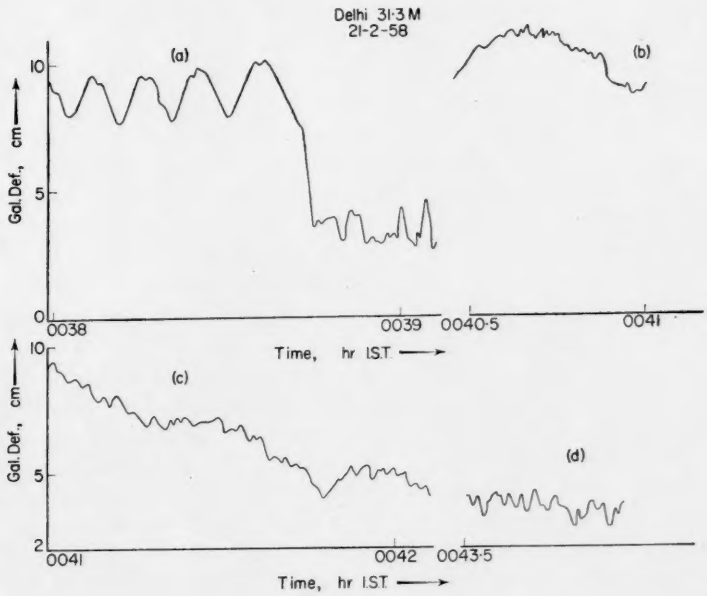

Delhi 31·3 M
21-2-58
(a)
(b)
(c)
(d)
Gal. Def., cm
10
5
0
2
0038
0039
0040·5
0041
0042
0043·5
Time, hr I.S.T

Fig. 2.

When the electronic density of the *F2*-layer is gradually increasing and the m.u.f. for the layer is approaching the frequency of the signal the sequence in which the magneto-ionic pattern would appear will be the reverse of what has been observed in Figs. 1 and 2. This reverse feature is perhaps shown by the fading curves of Fig. 3. The curves shown in this figure are for transmissions of wavelength 19·86 m from Delhi observed during 0611 to 0647 hours IST on 10 April 1958. During this period the ionic density of the *F2*-layer was likely to increase. Before 0634 hours the signal was found to be very weak and the nature of its fading was of a random type as shown in Fig. 3(a). This weak signal was obviously due to back scattering. At 0634 hours the intensity of the signal suddenly increased (Fig. 3b) perhaps due to the beginning of reflection of the extraordinary wave and was followed by a quick periodic fading which may be ascribed to interference between the upper and lower trajectory rays of the extraordinary wave. At 0635 hours there was a further sudden increase in the amplitude of the signal (Fig. 3c) which could be due to the onset of reflection of the ordinary wave. Soon after there appeared a fading pattern consisting of a quick periodicity superimposed upon a slower one (Fig. 3d). This may be due to the interference of the upper and lower trajectory rays of the ordinary wave and the lower trajectory ray of the extraordinary. In Figs. 3(e) and 3(f) one can see the increase of the frequency of the quicker periodicity and the decrease of that of the slower one which are as expected. Soon after 0646 hours the periodic pattern disappeared.

The fading curves of Figs. 4(a)–4(d) appear to be due to interference between waves singly and doubly reflected from the *F2*-layer. They are for a transmission of wavelength 31·48 m from Calcutta which is at a distance of about 450 km from the receiving station. During the period of observation any reflection from *E* or E_s-layer was not likely, but double reflection from the *F2*-layer was quite feasible. From such ionospheric data as were available for Calcutta we may infer that during the earlier part of the period of observation *F2*-layer could have a velocity of about 60 km/hr. This would give a frequency of 70 c/min for the periodic pattern formed by interference between waves singly and doubly reflected from the *F2*-layer and this agrees well with the observed frequency.

The fading curves of Figs. 4(e)1–4(g)2 were obtained at a time when there was no chance of any reflection from the *E*-layer and very little chance for double reflection from *F2* layer for the signals of wavelengths 16·86 m and 19·86 m transmitted from Delhi. At that part of the day when observations were taken, the ionic density of the *F2*-layer is generally on an increase but the disappearance of the signal of wavelength 16·86 m at about 1000 hours indicates that the ionic density of this layer had fallen at that time. Such a temporary fall is, however, quite feasible. If we assume that the ionic density of the *F2*-layer had started falling before 0957 hours and that it continued to do so till about 1008 hours when it again began to increase, all the features of the fading curves mentioned above may be explained. The sudden fall of the intensity of the signal of wavelength 16·86 m seen in Fig. 4(e)2 is very likely due to the disappearance of the extraordinary wave of the signal, and the periodicity preceding the disappearance (Figs. 4(e)1 and 4(e)2) may then be due to interference between the lower and upper trajectory rays of the extraordinary waves. When the signal of wavelength 19·86 m had penetrated the

VOL. 16 1959

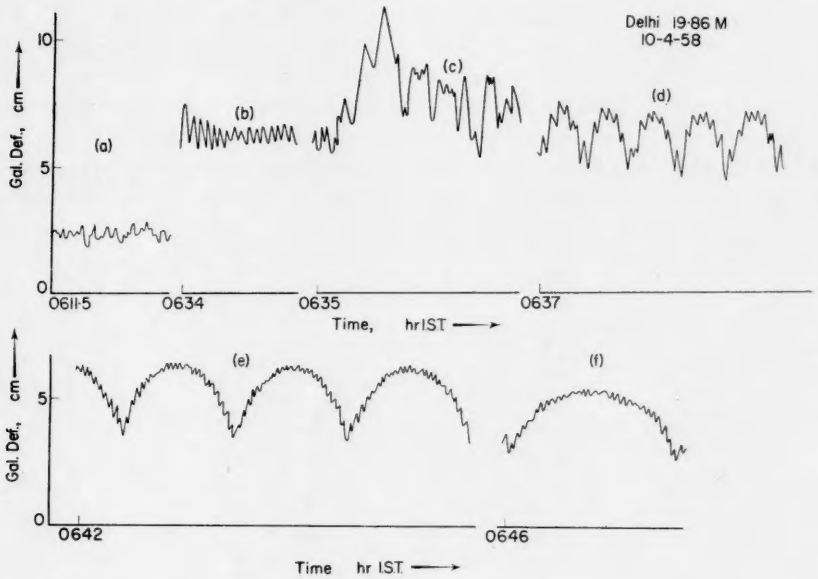

Delhi 19·86 M
10-4-58
(a)
(b)
(c)
(d)
Gal. Def., cm
10
5
0
0611·5
0634
0635
0637
Time, hr IST
(e)
(f)
0642
0646
Time hr IST

Fig. 3.

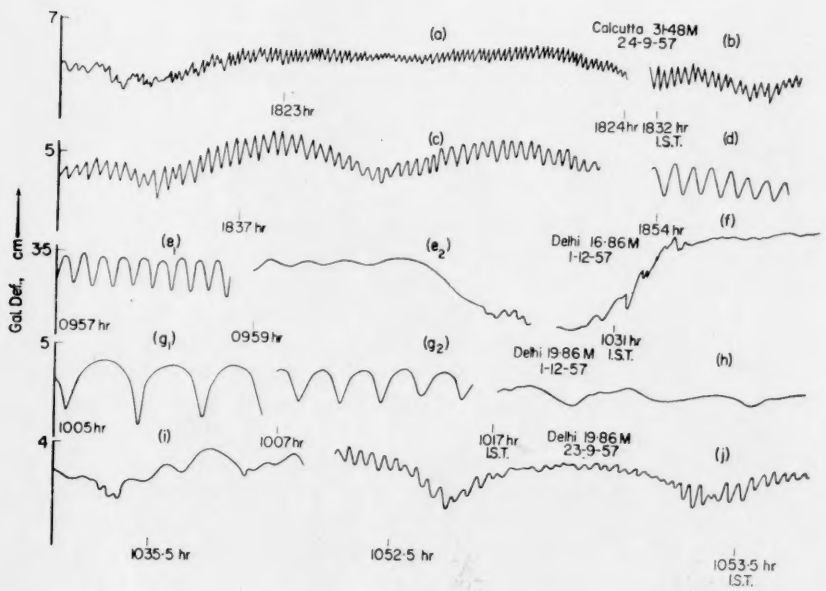

Calcutta 31·48M
24-9-57
(a)
(b)
(c)
(d)
(e1)
(e2)
(f)
(g1)
(g2)
(h)
(i)
(j)
1823 hr
1824 hr
1832 hr
I.S.T.
1837 hr
1854 hr
Delhi 16·86M
1-12-57
0957 hr
0959 hr
1031 hr
I.S.T.
Delhi 19·86 M
1-12-57
1005 hr
1007 hr
1017 hr
I.S.T.
Delhi 19·86 M
23-9-57
1035·5 hr
1052·5 hr
1053·5 hr
I.S.T.
Gal. Def., cm
7
5
3·5
5
4

Fig. 4.

F2-layer the signal of wavelength 19·86 m was still being reflected from this layer. The periodic patterns shown in Figs. 4(g)1–4(h) may be ascribed to interference between the upper and lower trajectory rays of the ordinary wave. The increase in the frequency of this periodicity up to about 1008 hours and the subsequent decrease (Fig. 4h) may be due to the fact that after 1008 hours the ionic density started to increase. That it became sufficiently great to cause reflection of the signal of wavelength 16·86 m is indicated by the curve in Fig. 4(f).

The fading curves of Figs. 4(i) and 4(j) are probably due to interference between waves simultaneously reflected from *E*- and *F2*-layers. At the time of observation

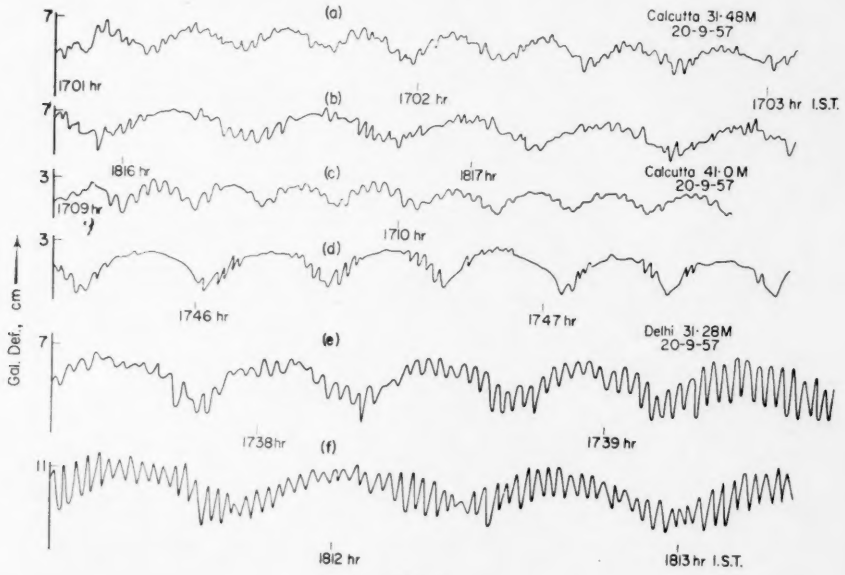


Fig. 5.

it was very likely that the ionic density of the *E*-layer had become just sufficient to cause reflection of the radio signal from this layer. At this stage interference is likely between the upper and lower trajectories of the extraordinary wave and the quick periodicity of the fading curves may be due to this. At this stage it was also likely that the radio signal reflected from the *F2*-layer was reaching the receiving station and the slow periodicity of the curves may be due to interference between this component and the lower trajectory of the extraordinary wave reflected from the *E*-layer.

The periodic type of fading curves of Figs. 5(a)–5(f) are for transmissions of wavelengths 31·48 m and 41·61 m from Calcutta and 31·3 m from Delhi. They are very much alike and their important common features are as noted below:

(a) All of them are of slow periodic type with a quick rhythmic periodicity of smaller amplitude superimposed upon the former.

(b) The periodic pattern has lasted for a very long time.

(c) The frequency of the slower periodicity increases with time. At a given

VOL. 16 1959

time, however, its value appears to be maximum for the signal of wavelength 41·61 m from Calcutta and minimum for the signal of 31·3 m from Delhi.

(d) The frequency of the quick ripples practically remains constant.

From the ionospheric data available for that day for Calcutta and Delhi it appears that while reflections of these wavelengths were not possible from the *E*-layer, the critical frequency of the *F2*-layer was sufficiently high to make the signals from Calcutta and Delhi reach the receiving station after suffering double reflection from the *F2*-layer. Even for these doubly reflected waves the signal frequencies for all the signals were sufficiently far off from the m.u.f. for the layer. The presence of ripples superimposed upon the slower frequency clearly indicates the existence of three interfering components, but it is difficult to explain the origin of these components in the present case. As the m.u.f. for the *F2*-layer, even for the doubly reflected wave, is sufficiently far from the frequency of any of the signals the fading curves are not likely of magneto-ionic origin. This is further corroborated by the fact the same type of pattern are being almost simultaneously obtained for signals of wavelengths 31·48 m and 41·61 m both transmitted from Calcutta. The quick fading may, however, be explained as due to interference between waves singly and doubly reflected from the *F2*-layer because the *F2*-layer is likely to have a vertical velocity at this time of the day. But difficulty remains regarding the origin of the slow fading.

During the course of the study of the periodic fading the authors have found several whose origin appears to be obscure.

References

Appleton E. V. and Beynon W. F. G. 1947 *Proc. Phys. Soc., Lond.* **59,** 58.
Banerjee S. S. and Mukherji G. C. 1946 *Nature, Lond.* **158,** 413.
Banerjee S. S. and Mukherji G. C. 1948 *Phil. Mag.*, **39,** 697.
Banerjee S. S. and Singh R. N. 1948 *Indian J. Phys.* **22,** 413.
Banerjee S. S. and Singh R. N. 1949 *Sci. & Cult.* **14,** 293.
Khastgir S. R. and Das P. M. 1950 *Proc. Phys. Soc., Lond.* **63,** 924.
Khastgir S. R. and Tantry B. A. P. 1951 *Indian J. Phys.* **25,** 217.
Singh B. N. and Ram R. L. 1958 *J. Atmosph. Terr. Phys.* In press.
Singh B. N. and Ram R. L. 1958 *J. Geophys. Res.* In press.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 156 to 159. Pergamon Press Ltd. Printed in Northern Ireland

The effect of trail irregularities on the interpretation of meteor echoes

A. G. McNAMARA and D. W. R. McKINLEY
National Research Council, Ottawa, Canada

(*Received* 15 *December* 1958)

Abstract—A brief discussion of some points raised in a recent paper by MANNING on obliquely-scattered meteor echoes is followed by the suggestion that the initial distribution of ionization along a typical meteor trail is markedly irregular. Several tentative hypotheses are advanced to account for the irregularities.

MANNING (1959) has endeavoured to explain some of the very knotty problems that arise from an examination of the observational evidence of forward and backward scatter from overdense meteor trails. We will conclude with a few general remarks on these difficulties, but we would like first to clarify some points of interpretation which have come up in MANNING's discussion of our observations. A conventional range-time display was used in the forward and backward scatter experiment of McKINLEY and McNAMARA (1956). It was therefore more convenient to measure the duration of the signal above the receiver noise threshold, than to attempt to obtain the echo decay-time directly. A meteor of transitional trail density, $q \simeq 10^{14}$ el/m, occurring in an average location in the observable region, should give a theoretical duration above threshold on our equipment of about one second as contrasted with the theoretical decay time-constant of 0·2 sec. This may be demonstrated by the following analysis in which we shall use the formula for a column of electrons scattering independently:

$$P_r = \frac{P_t G^2 \lambda^3 F^4 q^2 \sigma}{128\pi^3 R^3}$$

where $\sigma = 4\pi\left(\frac{e^2}{mc^2}\right)^2 = 1 \times 10^{-28}$ m^2 is the back scatter cross-section for one electron.

For our equipment, and a meteor trail occurring in an average location, the following parameters are applicable:

$P_t = 3 \times 10^5$ W.

$G = 24$, the maximum gain relative to an isotropic radiator.

$\lambda = 9$ m.

$F^4 = 0{\cdot}1$, the average correction to G for off-axis gain, taken from an inspection of Figs. 5 and 4 of our paper.

$q = 10^{14}$ el/m.

$R = 1{\cdot}5 \times 10^5$ m.

Substitution gives:

$P_r = 0{\cdot}94 \times 10^{-9}$ W.

VOL. 16 1959

The threshold sensitivity of the receiver is 1×10^{-14} W. Hence the signal/threshold ratio is $9{\cdot}4 \times 10^4$, or 50 dB. Assuming an exponential decay of amplitude due to a diffusion coefficient $D = 3$ m^2/sec, the decay time constant is

$$t_u = \frac{\lambda^2}{16\pi^2 D} = 0{\cdot}2 \text{ sec}$$

However, the time required for the initial amplitude A_0 to fall to the threshold level A is

$$T = t_u(2{\cdot}3)(\log A_0/A) = (0{\cdot}2)(2{\cdot}3)(2{\cdot}5) = 1{\cdot}2 \text{ sec}$$

for the transitional case.

Therefore, $T = 1{\cdot}2$ sec may be regarded as the minimum theoretical division point between the observed under and overdense trails, rather than $t_u = 0{\cdot}2$ sec. In practice, we adopted 3·5 sec for the transitional duration, since at this value a distinct change has been shown to take place in the character of the echoes (McKinley, 1953).

We have pointed out that it would be unlikely for the observed underdense trails to yield the full duration increase expected from using the value 2 for the exponent in the forward-scatter enhancement factor $\sec^m \phi$, both because of statistical contamination of the underdense group and because of secondary effects in the scattering mechanism. We have stated however, that our value of 1·73 for the exponent does indeed support the theoretical value of 2 for the limiting case of very low density trails.

In the long-enduring overdense cases, the meteor trails must eventually be strongly influenced by turbulence or distortion by winds. Moreover, any fragmentation will result in a rough trail capable of producing distributed targets. Therefore, it should not be surprising that simple geometrical models may fail to explain all the observed phenomena. In fact, as Manning has pointed out, on the theoretical model of a perfect cylindrical trail it should have been impossible for us to obtain forward duration enhancements greater than that corresponding to an exponent of 0·4, whereas our experimental average was 1·13. If it were assumed that random trail distortion would eventually make all orientations equally probable, then the mean exponent of the cylindrical configurations could well lie in the region of our value.

However, if trail distortion occurs during the echo lifetime, it is quite likely, on the distorted smooth cylinder model, that the reflection points for the backscatter and the forward-scatter are located on sections of the trail that have orientations (ϕ and β of Manning, 1957) which are more or less uncorrelated with each other and with the original trail parameters. On these assumptions the longer the trail duration the poorer will be the correlation. A meteor occurring at a given time would yield one value of the forward enhancement factor, but at some other time a different value of the factor would be deduced from the identical meteor, owing to the change in wind pattern. The only meaningful value for the enhancement will be an average over a large number of measurements. In particular, it might appear that Manning's overdense theory can be applied accurately to a single observation only if it falls in a very small range of trail durations. That is, the trail must have $q > 10^{14}$, and at the same time the duration should be shorter

than that for which trail distortion becomes predominant, which according to Greenhow (1952) should occur in the range 0·2–1·6 sec.

It might be useful to repeat the experiment of McKinley and McNamara with the addition of a back-scatter radar at the middle of the path, beamed vertically. For this geometry, meteor trails having all β-values from 0° to 90° should be observable simultaneously by back-scatter and forward-scatter, and a significant variation in the exponent should be observable. In fact, if radiant information is available, it should be possible to examine the extremes of trail orientation separately and to test for the effect of β.

Manning's theory of the loss of aspect sensitivity (Manning, 1957), based on the distortion by winds of originally straight and uniformly ionized trails, has been shown to fit some of the observations. However, it fails to account for the not-uncommon instances where the enduring ionization appears at points on the path far removed from t_0 instantly after the passage of the meteor (within 0·02 sec at least, as seen on the radar records). On the other hand, the hypothesis of a rough or irregularly-ionized trail can be developed to explain all these observations in general terms. The observed increase in delay of appearance of certain enduring echoes with distance away from the t_0 point may be explained either by Manning's distortion theory, which should be applicable also to the rough trail, or by the diffusion mechanism proposed by McKinley and Millman (1949), or by a combination of both. For this purpose it is assumed that the irregularities are created as segments many metres in length; at distances remote from t_0 each segment will include several Fresnel zones, resulting in cancellation of the scattered components until distortion or diffusion takes place. In other cases, the instantaneous appearance of enduring ionization, also far from t_0 and often with a continuum of ranges, could be due to random variations in ionization occurring on a shorter scale, such that initial cancellation of the Fresnel zone contributions does not occur or is not complete. The difference between the rough and smooth trail hypotheses becomes less apparent near t_0, where the Fresnel zones are many hundreds of metres in length, and the radio wave is scattered substantially in phase from all the irregularities occurring along one zone segment. As Manning points out, the idealized smooth trail alone is very difficult to treat theoretically, and the detailed analysis of the problem becomes vastly more complicated when random irregularities are introduced.

It would seem from the observational evidence that the initially smooth meteor trail may be the exception, and that irregularity and disorder may be the rule. Winds alone cannot explain all the phenomena, and one must look for other factors. The variations in ionization along the trail might be produced by minor fragmentation effects, by stratification in the ionosphere, or through rapid rotation of an irregularly-shaped meteoroid. Rotation might be expected to cause a regularly-repeated variation, which is indeed suggested by some of the observations. Ionization produced by ultra-violet radiation might also be a contributory factor. Chemical reaction between some of the lighter elements of the meteor and the atmospheric particles could provide additional energy for the creation of ionization. There could well be variations in the relative rates at which different components of a meteoroid were vaporized, owing to the heterogeneous character of the body,

VOL. 16 1959

whether or not the object were rotating. Small-scale turbulence, if it exists in the atmosphere at meteoric heights, should be capable of roughening an initially smooth trail eventually, but the velocity of turbulence would have to be very great to account for echoes appearing in less than 0·02 sec. We would hope that some of the suggestions offered here to account for the rough meteor trail may be regarded as fairly reasonable, but we are aware that none of them is likely to ease the theoretician's already difficult task.

References

Greenhow J. S.	1952	*Proc. Phys. Soc.* B. **65,** 169.
McKinley D. W. R.	1953	*Canad. J. Phys.* **31,** 1121.
McKinley D. W. R. and McNamara A. G.	1956	*Canad. J. Phys.* **34,** 625.
McKinley D. W. R. and Millman P. M.	1949	*Proc. Inst. Radio Engrs. N. Y.* **37,** 364.
Manning L. A.	1957	*J. Atmosph. Terr. Phys.* **15,** 137.
Manning L. A.	1959	*J. Atmosph. Terr. Phys* **14,** 82.

OL.
16
959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 160 to 173. Pergamon Press, Ltd. Printed in Northern Ireland

A new theoretical model of the composite *F*-layer

F. MARIANI
Istituto di Fisica dell' Universitè, Roma

(*Received* 21 *September* 1957; *revised* 10 *November* 1958)

Abstract—A model of a composite *F*-layer is developed to explain both the regular behaviour of the *F1*-layer and the anomalies of the *F2*-layer. The bifurcation in the *F1*- and *F2*- layers arises as an effect of a discontinuity in the physical conditions of the upper atmosphere, namely a fairly rapid variation of the temperature vertical gradient. The assumed height variation of the generalized recombination coefficient is $\alpha = \alpha_0 n^{\xi} N^{-1} T^{-1}$. Tidal movements have a smaller importance, because they cause only secondary perturbative effects.

1. INTRODUCTION

RECENT investigations on the ionospheric *F*-layer seem to support the conception of an uninterrupted height distribution of electron density, the distinction between the *F1*- and *F2*-layers not being sufficiently clear-cut to allow them to be considered as separate layers.

In our opinion, we can think in terms of some more or less sudden variation of any parameter characterizing the physical conditions of a single layer, to which corresponds a more-or-less sudden variation in the vertical gradient of electron density. The two layers *F1* and *F2* would, rather, seem to be the upper and lower parts of a single layer. If one accepts this point of view, it is possible to study, on a new basis, the model of a single *F*-layer first suggested by BRADBURY (1938).

Recently, some authors have considered various possibilities: for example MITRA (1952) and RATCLIFFE (1956) have assumed certain different explicit mathematical forms of the loss coefficient at various heights, suggested by physical considerations; however, their results must to some extent be revised, because they have assumed a simple Chapman's law for the ion production rate: in other words, they do not consider the effects of the height variation of the temperature. On the other hand, the experimental results, on which the recent extension of the standard atmosphere is based, seem to support the view of some more or less sudden variation of the vertical temperature gradient occurring at about 180 to 200 km above the ground. The results of HAVENS *et al.* (1954) and those of KALLMANN *et al.* (1956) seem to agree on the occurrence of a fairly sudden variation of the temperature gradient at from 220 to 280 km above the ground.

The height of the observed temperature gradient variations could be identified with the geometrical height of the *F1*-layer, which is usually about 200 to 250 km, as JACKSON (1956) and other authors have also recently shown. Moreover the method of analysis worked out by JACKSON, applied to various cases of ionograms obtained for higher and lower solar activity, has made it possible to ascertain a remarkable degree of regularity in the day-time heights of the *F1*- and *F2*-layers, notwithstanding the variations in the electron density occurring during the sunspot cycle.

It is the purpose of this paper to study the effects of the preceding suggestions on electron density in the *F*-layer; we calculate the effective theoretical, diurnal

VOL. 16 1959

and seasonal variations of the critical frequencies and heights. We assume, as BATES (1954) did, a temperature increasing linearly up to a certain height, about 250 km above the ground, and a constant temperature for higher levels.

2. General Considerations on the Electron-loss Equation and Drift Velocity

We assume that the height and time variations of electron density are governed, at a particular level, by the equation

$$\frac{\partial N}{\partial t} = I(z,\chi) - \alpha N^2 + R_T + R_M \tag{1}$$

where N is the electron density at height z; $I(z,\chi)$ is the ion production rate, depending upon z and the zenithal angle χ of the sun; α is the generalized recombination coefficient given by BATES and MASSEY (1946), not including the part explicitly depending upon the time variation of the temperature, which we have included in the term R_T; R_M is a term that includes the effects due to the movements of the ionosphere, for example, those due to tides.

Concerning the height dependence of the temperature $T(z)$, we write

$$T(z) = \begin{cases} T_0(1 + pz) & \text{if } z \leqslant c \\ T_1 = T_0(1 + pc) = \text{a constant} & \text{if } z \geqslant c \end{cases} \tag{2}$$

where we have assumed $z = 0$ at the base of the F-layer; $T_0 = T(0)$; p is a constant and c is a reference level, located at about 200 to 250 km above the ground.

The values of T_0, and of the vertical gradient $T_0 p$ of the temperature T, can *a priori* depend upon the time; as regards this question, one can assume, at least during daylight (except during twilight) that the time variation of T_0 and p is small. We have already considered this problem (MARIANI, 1956c) and we have shown with good evidence that one may exclude the possibility of large diurnal and seasonal temperature variations. These results agree with those that may be deduced for a consideration of radiative equilibrium (MARTYN and PULLEY, 1936) in the atmosphere, according to which the effective variations are no larger than from 20 to 25 per cent. As a general conclusion, we have asserted that the large values for the daylight variations of the temperature, recently assumed by some authors in order to explain the anomalies of the $F2$-layer, are probably excessive and, furthermore, do not always give good results. Concerning the night and the twilight variations, we shall later assume that, also under these conditions, the values of T_0 and p are constant. Finally, also the height c is assumed independent of time.

The explicit form of the ion production rate $I(z,\chi)$ for the temperature distribution (2) has already been studied by the author (MARIANI, 1955, 1956a) for cases of both finite and infinite radius of the earth. On the basis of the preceding hypotheses, the term R_T may be neglected; instead, the term R_M, the effects of which MARTYN (1947) was the first to study extensively, has the mathematical

expression $-\partial(Nv)/(\partial z)$ if we assume a plane stratification of the ionosphere and indicate by v the vertical drift velocity of the electrons, that may even depend on the height z.

If one considers, first of all, α as a constant, the equation (1) assumes the form

$$\sigma \frac{\partial \nu}{\partial \Phi} = F - \nu^2 - 1{\cdot}37 \times 10^4 \, \sigma \frac{\partial v}{\partial z} \nu - 1{\cdot}37 \times 10^4 \, \sigma v \frac{\partial \nu}{\partial z} \tag{3}$$

where, with the usual notation

$$\begin{aligned} \Phi &= 2\pi t/86{,}400 = t/(1{\cdot}37 \times 10^4) \\ F &= I/I_0 \\ I_0 &= \text{the maximum of } I(z,\chi) \text{ for } \chi = 0^\circ \\ \nu &= N/N_0 \\ N_0 &= (I_0/\alpha)^{1/2} \\ \sigma &= 1/[1{\cdot}37 \times 10^4 (N_0 \alpha)] \end{aligned}$$

A first observation concerns the values of σ: if one accepts the values of N_0 and α assumed by BATES and MASSEY, these are $\sigma = 1/13{\cdot}7$ for the *F1*-layer and $\sigma \simeq 0{\cdot}9$ for the *F2*-layer.

We have shown (MARIANI, 1956a) that, neglecting the term R_M, the calculated variation of f_0F1 at 45° of geographical latitude fits the experimental values better if one assumes $\sigma = 1/25$ rather than $\sigma = 1/5$. These rather low values have an order of magnitude equal to the values considered by BATES and MASSEY, so that the neglect of R_M in the *F1*-layer is justified *a posteriori*.

Concerning the *F2*-layer, the observed daily excursion of f_0F2 is normally larger in winter than in summer, so that we can infer some seasonal variation of σ: the value of σ in summer would be larger than in winter. This point of view agrees, for example, with the results of MITRA (1951) who has given the values $\alpha = 3 \times 10^{-11}$ and $\alpha = 1{\cdot}5 \times 10^{-11}$ cm^3/sec in the *F2*-layer, respectively for winter and summer conditions.

At this point, we wish to ascertain when in the *F2*-layer the terms $-1{\cdot}37 \times 10^4 \, \sigma v(\partial v)/(\partial z)$ and $-1{\cdot}37 \times 10^4 \, \sigma v(\partial \nu)/(\partial z)$ become important with respect to the "static" terms F and $-\nu^2$. Let us consider a summer day at a medium latitude; if, according to the experimental results, we wish to obtain from equation (3) values of the maximum (reduced) electron density ν_m increasing towards sunset, we have to assume positive values for the right-hand side of equation (3), at least at some level near that of maximum electron density. In other words, the term R_M has to be larger than the term ν^2, especially around sunset, when F is vanishing. At first, we assume a drift velocity v independent upon the height z and in the correct phase-angle in order to make the term R_M positive. In effect the term $1{\cdot}37 \times 10^4 \, \sigma v(\partial \nu)/(\partial z)$ is comparable with ν^2 only if we assume rather large values of v. If we put $\sigma = 1$, $\nu = 1$, this gives:

$$v = 1 \Big/ \left(1{\cdot}37 \times 10^4 \frac{\partial \nu}{\partial z}\right) = H \Big/ \left(1{\cdot}37 \times 10^4 \frac{\partial \nu}{\partial Z}\right)$$

VOL. 16 1959

where H is the scale height and Z is the reduced height. Near the height corresponding to the maximum of ν we can reasonably assume $H \simeq 60$ km and $1/\nu\,(\partial\nu)/(\partial Z) \leqslant 10$ per cent, so that the resulting actual value of v is of the order of 50 m/sec.

We now consider the case $v = v(Z)$ and assume $v = V(Z)\,f(\Phi) = v_0 \exp(-\gamma Z)\,f(\Phi)$, where $f(\Phi)$ is a sinusoidal term and γ a constant of the order of unity. In this case, the principal part of R_M is expressed by the term $1{\cdot}37 \times 10^4 \sigma\nu(\partial v)/(\partial z)$, so that $v_0 = H \exp(\gamma Z)/[1{\cdot}37 \times 10^4 \gamma\sigma\nu f(\Phi)]$; if we put $\gamma = 0{\cdot}5$, $\nu = 1$, $\sigma = 1$, $Z = 2$ and $H = 60$ km we have $v_0 \geqslant 24$ m/sec.

We conclude that the term R_M has the desired order of magnitude around summer sunset, only if the amplitude v_0 of the drift velocity is not less than about 24 m/sec. This last value might seem too large. If we consider a sinusoidal velocity $v = v_0 \sin \omega t$, the corresponding height excursion Δh during a period T is $\Delta h = v_0 T/\pi$; with $T = 12$ hr and $v_0 = 25$ m/sec, we have $\Delta h \simeq 340$ km; but when $T = 24$ hr this becomes $\Delta h \simeq 680$ km. If we consider, instead, a velocity $v = v_0 \exp(-\gamma Z) \sin \omega t$ the excursion Δh corresponding to a level $Z = Z_0$ for $t = 0$ is:

$$\Delta h = H\left[\frac{1}{\gamma}\left\{\log \exp(\gamma Z_0) + T\frac{\gamma v_0}{H\pi}\right\} - Z_0\right]$$

We remember that the geometrical height z corresponding to $h'F2$ is always larger than the height c, so that we can calculate the value of the excursion Δh corresponding to the level $Z_0 = 0{\cdot}5$; with $\gamma = 0{\cdot}5$, $H = 60$ km, $v_0 = 25$ m/sec we have $\Delta h \simeq 140$ km and $\Delta h \simeq 200$ km respectively for $T = 12$ hr and $T = 24$ hr. In every case, the calculated values Δh, which are very likely only lower limits, seem too large for various reasons. First of all, they represent geometrical height excursions to which should correspond considerably larger virtual height excursions $\Delta h'F2$ of the $F2$-layer. The normal daily effective excursion of $h'F2$, including the daily excursion of the level of maximum ion production, is usually of the order of 100 km. In effect, SATO and NAMIKAWA (1954), in their study on latitudinal distribution of diurnal and semi-diurnal components of $h'F2$ in thirty-six observatories, have obtained values of the diurnal excursion not exceeding 100 km, at higher northern latitudes, but rather smaller at southern and lower northern latitudes. We may assume, then, that the effective geometrical height excursion due to the drift does not exceed some tens of kilometres; the drift velocities corresponding to the effective Δh thus seem considerably smaller than the minimum values we are forced to assume in order to explain the summer diurnal anomaly of f_0F2.

Moreover, we note that in winter, when the $F1$-layer does not appear, the minimum virtual height $h'F2$ generally reaches a minimum during the meridian hours, in agreement with theoretical forecasts. Thus, some doubt arises on the physical meaning of the high summer values of $h'F2$ and of its meridian maximum, which could be due to some greater retardation than during the winter. Such a point of view also seems suggested by a comparative study of the daily behaviour of $h'F2$, f_0F2 and f_0F1. This study shows some correlation of $h'F2$ with the difference $(f_0F2 - f_0F1)$. For this reason we can thus reasonably think in terms

of considerable group retardation effects. These effects are produced by the *F1*-layer and, at least in many cases, may mask or alter the effective daily excursion of the height of the *F2*-layer. Regarding this question, we have shown (DOMINICI and MARIANI, 1955, 1956) in some particular cases the possibility that larger meridian values of $h'F2$ may correspond to a diminution of the true geometrical height.

The above considerations raise some doubt on the importance of the required values of the vertical (not the horizontal!) drift velocities and on their phase angle, as inferred from the analysis of the daily variation of $h'F2$.

Many authors have, however, studied the behaviour of the *F2*-layer, and have assumed suitable values of the drift velocities; sometimes their assumptions have been based on theoretical considerations of the electrical conductivity of the upper atmosphere. Many aspects of the anomalies of the *F2*-layer have been explained; however, the values of the drift velocity sometimes lead to excessive daily height excursion: for example HIRONO and MAEDA (1954, 1955) have obtained daily excursions from 150 to 200, even up to 300 km, which in our opinion cannot be accepted without reserve. On the other hand, SATO has shown (1954) that the seasonal anomaly, as also the different increase of the electron density from sunspot minimum to sunspot maximum, which is larger in winter than in summer, cannot be explained only by the assumption of vertical drift.

VOL. 16 1959

3. THE NEW THEORETICAL MODEL OF THE COMPOSITE *F*-LAYER

Our model of the *F*-layer is based on a simple, but physically significant, analytical expression of the generalized recombination coefficient α; we wish to ascertain the possibility of accounting for the diurnal and seasonal variations of the maximum electron density and height of both the *F1*- and *F2*-layers.

A fairly general expression of α, which was tentatively and advantageously assumed by BURKARD (1956) for the *F1*-layer, is

$$\alpha = \alpha_0 n^{\xi} N^{n} T^{\zeta} \tag{4}$$

where n, N, T are respectively the particle concentration, the electron density and the absolute temperature; α_0, ξ, η, ζ are constants. When $\xi = 1$, $\eta = \zeta = -1$ we obtain the expression $\alpha = \alpha_0 n (NT)^{-1}$ which, by certain particular hypotheses, BARAL and MITRA (1950) have derived theoretically. From now on, we shall assume $\eta = \zeta = -1$ and $\xi =$ a positive value probably not larger than unity.

It must now be remembered that, according to current opinion, the electron loss equation has the form of a recombination equation for the *F1*-layer and an attachment equation for the *F2*-layer. In other words, the second term on the right hand side of equation (1) is respectively quadratic or linear with respect to the electron density N. On the other hand, the diurnal variation of the critical frequency f_0F1 agrees well with the assumption $\alpha = \alpha_c =$ a constant at the height of maximum *F1* electron density.

As follows from this last condition and from equation (4), the coefficient α may be written in the form

$$\alpha = \alpha_c (n/n_c)^{\xi} / [(N/N_c)(T/T_c)] = \alpha_c \phi \nu_c / \nu \tag{5}$$

where α_c is the constant value of α at the height c; n_c, N_c, T_c are respectively the values of n, N, T at the same height; ϕ is a known function of n and T.

If we put (MARIANI, 1956a):

$$u = \frac{1 + pz}{1 + pz^*}$$

with z^* the height of maximum ion production rate when $\chi = 0°$

$$s = \frac{1 + pc}{1 + pz^*}$$

$$Z = \frac{z - c}{(1 + pc)H_0}$$

where H_0 is the scale height at $z = 0$

$$r = -1/(pH_0)$$

$$\mu = 1 + pH_0$$

we obtain

$$\phi = \begin{cases} (s/u)^{1+\xi(1-r)} & \text{for} \quad z \leqslant c \\ \exp(-\xi Z) & \text{for} \quad z \geqslant c \end{cases} \tag{6}$$

and

$$F = \begin{cases} \exp\left[\mu(1 - \sec\chi \,.\, u^r)\right] \,.\, u^{r-1} & \text{for} \quad z \leqslant c \\ \exp\left\{\mu[1 - \sec\chi \,.\, s^r \,.\, \exp(-Z)] - Z\right\} \,.\, s^{r-1} & \text{for} \quad z \geqslant c \end{cases} \tag{7}$$

If we put equation (5) into equation (1), this may be written, with the usual notation:

$$\sigma \frac{\partial v}{\partial \Phi} = F - \phi v_c v - 1{\cdot}37 \times 10^4 \sigma v \frac{\partial v}{\partial z} - 1{\cdot}37 \times 10^4 \sigma v \frac{\partial v}{\partial z} \tag{8}$$

We note that $\sigma = 1/(1{\cdot}37 \times 10^4 N_0 \alpha_c)$ is a constant; conversely, the function ϕv_c depends upon height and time. However, the physical meaning of ϕv_c seems to suggest as a "natural" condition the existence of some lower limiting value k.

An interesting feature of equation (8) is that, if at height c we neglect the term R_M, it appears quadratic with respect to the variable v. At a height rather greater than c, around the height of maximum electron density in the $F2$-layer (which in our scheme is the upper part of the composite F-layer), the equation (8) appears substantially linear with respect to v, because at that height the term containing $(\partial v)/(\partial z)$ is negligible.

Thus the assumed expression (5) of α seems to account for the mathematical expression of the fundamental equation, especially if we consider that our knowledge on the quadratic or linear dependence on N in the equation (1) derives substantially from a study of the critical frequencies, i.e. from the study of N at height c and at height $z > c$, near its maximum value.

Later, we shall consider typical solstitial days at 45° geographical latitude and equinoctial days at different latitudes. Figs. 1 and 2 give the hourly median values of f_0F2 (and, when possible, of f_0F1) for the solstitial months of December 1953 and June 1954 (period of minimum solar activity), and, respectively, for December 1955 and June 1956. The data given in Fig. 1 refer to comparable

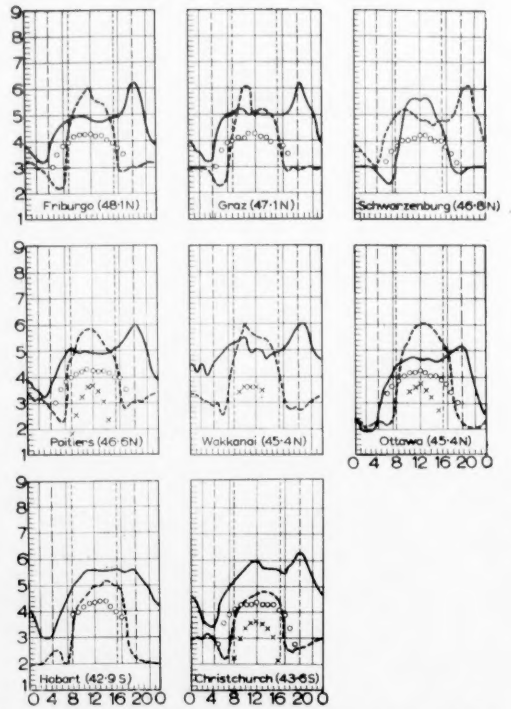


Fig. 1. Critical frequencies in December 1953 and June 1954:

---- and × × ×, f_0F2 and f_0F1 in winter;

—— and ○ ○ ○, f_0F2 and f_0F1 in summer.

Frequencies in MHz.

The numbers in parentheses are the geographical latitudes. The vertical marks indicate summer (— — —) and winter (– – – –) sunrise and sunset.

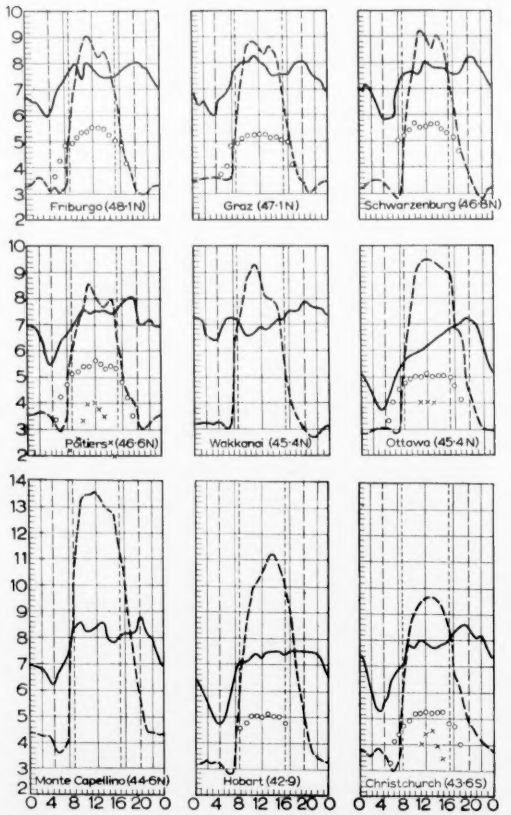


Fig. 2. Critical frequencies in December 1955 and June 1956 (at Monte Capellino, June and December 1956; at Schwarzenburg, December 1955 and July 1956). Symbols as in Fig. 1.

VOL. 16 1959

ionospheric conditions, corresponding to approximately equal solar activity; as regards Fig. 2, the level of solar activity is very different in the two consecutive solstitial months, so that the data for June give higher values than for December 1955.

The known diurnal and seasonal anomalies of f_0F2 appear clearly in the northern hemisphere. In the southern observatories of Hobart and Christchurch, on the

Table 1

		s	σ	ξ	k
a	Common values	1·582	1/5	0·9	0·05
b	$r = -2$	1·582	1/10	0·9	0·05
c	$1 + pz^* = 2$	1·582	1/10	0·9	0·025
d	$p = 0{\cdot}025$ km^{-1}	1·582	1/25	0·9	0·05
e		1·582	1/10	0·7	0·05
f		1·582	1/10	0·7	0·025
g		0·895	1/10	0·9	0·05

other hand, the seasonal anomaly is not present, at minimum solar activity. A remarkable feature of the winter night behaviour of f_0F2 is the almost systematic occurrence of a secondary maximum which, however, does not appear during summer nights. This fact suggests some semidiurnal tidal effect at the highest ionospheric levels which is effective (or in the correct phase, or comparable with the ground level of electron density) only, or mostly, during the winter. Finally,

Table 2

		k	$1{\cdot}37\times10^4$ v_0/H	T (hr)	
h	Common values	0·05	1	24	dotted lines in Fig. 8
	$r = -2$; $1 + pz^* = 2$	0·025	2	12	full lines in Fig. 8
	$s = 1{\cdot}582$; $p = 0{\cdot}025$ km^{-1}				
i	$\sigma = 1/10$; $\xi = 0{\cdot}7$	0·05	2	24	dotted lines in Fig. 8
	$\gamma = 0{\cdot}05$; $\delta = 180°$	0·025	1	12	full lines in Fig. 8

one notes the occurrence of two winter diurnal maxima of f_0F2 in some northern observatories, especially at high solar activity.

In order to test our model of a composite F-layer we have solved the equation (8), with ϕ and F given by equations (6) and (7), for some particular values of the parameters involved, indicated in the Tables 1 and 2, for solstitial days at 45° geographical latitude (tides absent or present, respectively).

Case of the F2-layer

The calculated values of $\nu_m^{1/2}$ and $\nu_c^{1/2}$, respectively proportional to the critical frequencies of the $F2$- and $F1$-layers, and the height Z_m of the maximum of electron

density for the various cases are shown in Figs. 3 and 8. First, let us consider the case when tides are absent (Fig. 3). If we assume $\xi = 0{\cdot}9$, two summer maxima of $\nu_m^{1/2}$ occur: an evening maximum around sunset and a smaller morning maximum, occurring later for increasing values of σ and occurring from 3 to 5 hr before noon. A meridian depression occurs around 2 p.m. Conversely, in winter, a single maximum occurs, rather late in the afternoon. With the exception of the case $\sigma = 1/5$,

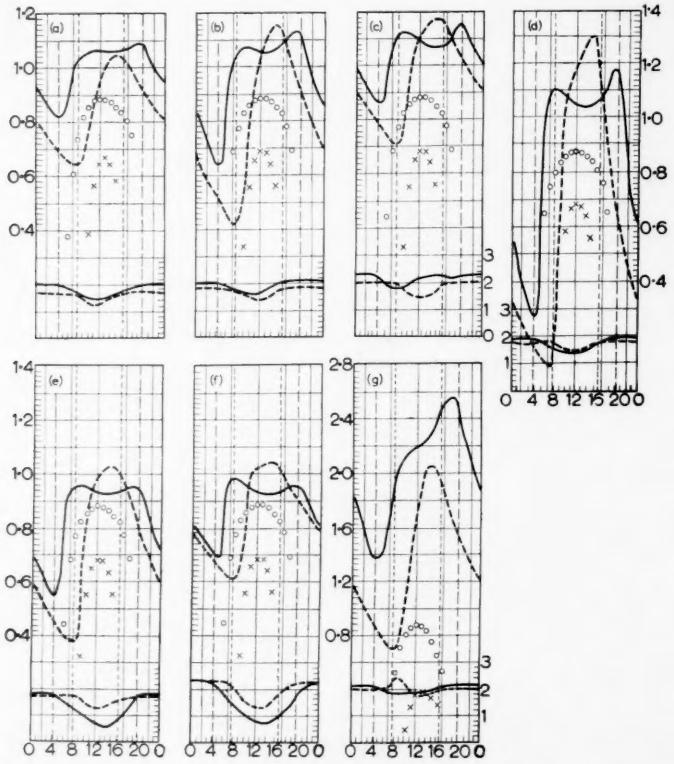


Fig. 3. Calculated behaviour of $\nu_m^{1/2}$ and $\nu_e^{1/2}$ (upper graphs and points) and of Z_m (lower graphs); tides absent.
——and ○ ○ ○, summer values; - - - - and × × ×, winter values.
Sunrise and sunset indicated as in Fig. 1.

the winter maximum is a little larger than the summer absolute maximum. Concerning the night behaviour, the summer values of $\nu_m^{1/2}$ are systematically larger than the winter values.

The only difference between the cases of different values of k is in the amplitude of the diurnal and seasonal excursions of $\nu_m^{1/2}$.

With regard to the heights, one cannot compare Z_m directly with the experimental values of $h'F2$, which are minimum virtual heights. However, the limited range of its diurnal and seasonal excursions is apparent. The diurnal behaviour is rather different for different values of k: if $k = 0{\cdot}05$, summer and winter meridian minima appear, while if $k = 0{\cdot}025$, Z_m reaches maxima and minima approximately corresponding to minima and maxima of $\nu_m^{1/2}$. This latter

result is similar to that obtained in a preceding paper (MARIANI, 1956c) corresponding to the cases (a) and (d) in Table 1 but without any lower limit of $\phi\nu_e$; Fig. 4 shows this very behaviour of ν_m and Z_m. However, in Fig. 3 the diurnal and seasonal excursion of ν_m have the correct order of magnitude and, moreover, the seasonal anomaly is quite satisfactorily explained.

Fig. 5 shows the time variation of the ratio α/α_c at the height of the maximum electron density of the $F2$-layer for case (b) of Fig. 3.

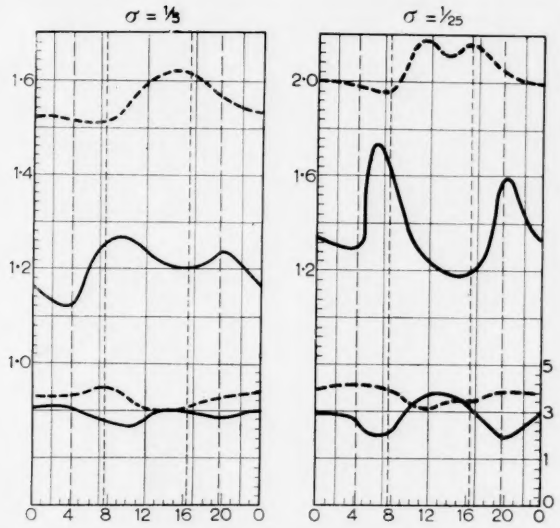


Fig. 4. Full and dotted lines indicate respectively summer and winter values of ν_m and Z_m.

With regard to the effects of different values of the parameter s, which regulates the height of maximum ion production with respect to the height c of the gradient discontinuity, we can see that the only appreciable differences between the cases $s = 1{\cdot}582$ and $s = 0{\cdot}895$ are the larger numerical values of ${\nu_m}^{1/2}$ and the systematic occurrence of electron densities that are higher in summer than in winter, in the case $s = 0{\cdot}895$. Concerning Z_m an unexpected winter maximum appears rather early in the morning.

We have also studied the noon latitudinal dependence of ${\nu_m}^{1/2}$ in an equinoctial day for the two cases $k = 0{\cdot}05$ and $k = 0{\cdot}10$, in correspondence to case (b) of Table 1. A maximum appears at a fairly high latitude, with a tendency to occur at a lower latitude for increasing values of k (Fig. 6).

Fig. 7 shows the effective height distribution of ν corresponding to four values of the latitude λ, which in our hypotheses (equinoctial day at noon) coincide with the zenital angle χ.

The latitudinal behaviour of ${\nu_m}^{1/2}$ is qualitatively correct.

We have also studied the effect of varying the value of ξ, assuming $\xi = 0{\cdot}7$ (and $\sigma = 1/10$, with $k = 0{\cdot}05$ and $k = 0{\cdot}025$). The main difference from the preceding results is that the morning summer maximum is larger than the afternoon maximum.

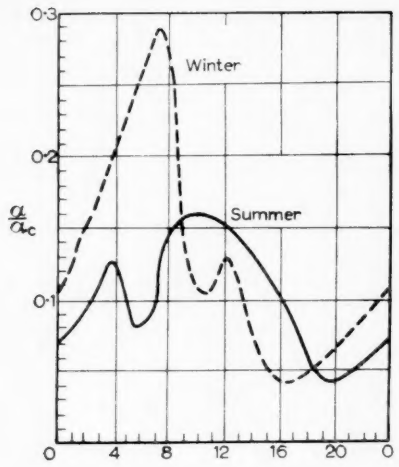


Fig. 5.

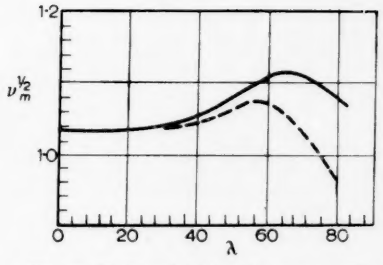


Fig. 6. Full and dotted lines correspond respectively to $k = 0{\cdot}05$ and $k = 0{\cdot}10$.

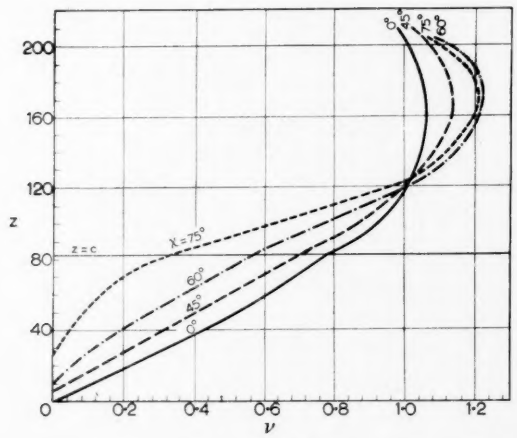


Fig. 7.

In conclusion, the general calculated behaviour of $\nu_m^{1/2}$ in the absence of tides may be considered to be in fairly close agreement with the experimental results, with the exception of the hour of occurrence of the winter diurnal maximum and of the secondary midnight maximum, which is systematically absent. In our opinion, these exceptions may be accounted for if we consider the secondary effects produced by some other cause, particularly by tidal, or more general, movements of the ionosphere.

We have, therefore, studied the modifications produced in $\nu_m^{1/2}$ and Z_m, for the winter solstice and for $\xi = 0{\cdot}7$, $\sigma = 1/10$, by the assumption of a vertical drift velocity $v = v_0 \exp(-\gamma Z) \operatorname{sen}(\omega\Phi + \delta)$; the values of the constants are also given in Table 2. We have considered the cases of both semidiurnal and diurnal periodicity. The values of v_0 for $H = 60$ km are respectively 8·8 and 4·4 m/sec

VOL. 16 1959

for semidiurnal and diurnal velocity; the corresponding excursion Δh at the level $Z_0 = 0{\cdot}5$ is about 70 km. Although the exponential term $\exp(-\gamma Z)$ produces a rapid reduction of the velocity amplitude for an increase in Z, the effects on $\nu_m^{1/2}$ (Fig. 8) are fairly considerable.

The time of diurnal maximum of $\nu_m^{1/2}$ shifts considerably towards noon; moreover, a small secondary night maximum of $\nu_m^{1/2}$ may occur before midnight.

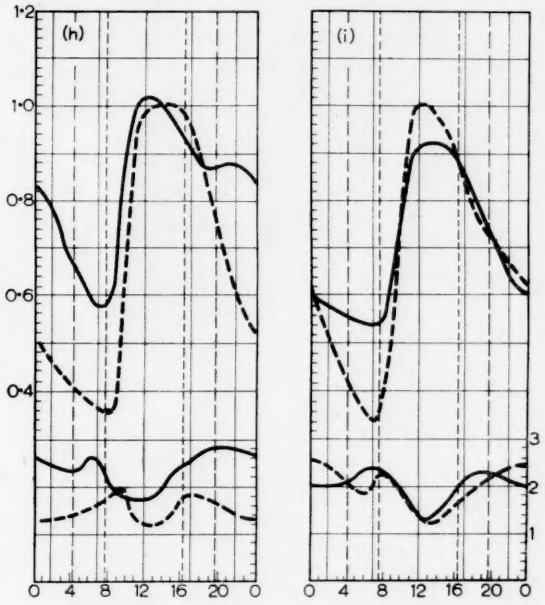


Fig. 8. Calculated behaviours of $\nu_m^{1/2}$ (upper graphs) and of Z_m (lower graphs); tides present.

Naturally, the numerical results may differ to a greater or lesser extent, if we assume different values of the parameters involved; the most interesting question is the possibility of obtaining a theoretical scheme which can account for all the experimental features of the $F2$-layer, by the assumption of only small vertical drift velocities.

Case of the F1-layer

The maximum electron density of this layer, according to our hypotheses, occurs at a constant height c. The values of $\nu_c^{1/2}$, calculated by neglecting the term R_M, agree with the experimental values; the tidal term R_M is very probably rather small as compared with the other terms of the equation (8), on account of the rapid increase of the factor ϕ in the term $\phi\nu_c\nu$ at lower heights.

The minimum height of the layer reaches a minimum in meridian hours, when the zenithal angle χ is smaller, and increases toward twilight. When χ increases, the discontinuity of the electron density gradient at the height $z = c$ becomes more and more negligible. These features, which can be directly derived from the equation (8) by assuming stationary conditions, appear clearly also in Fig. 7 which gives the noon latitudinal variation of $\nu(z)$ in an equinoctial day.

4. DISCUSSION AND CONCLUSIONS

Our results are, to some extent, qualitative rather than quantitative, because the values assumed for the theoretical parameters are more or less arbitrary. However, the general agreement of the theoretical with the experimental morphology of both the *F1*- and the *F2*-layers, by assuming only small corrective tidal terms, is encouraging.

Naturally, a complete theory of the layer should also consider the other secondary effects due to diurnal, seasonal or geographical temperature variations, to diffusion [which, if present, is more relevant in the early morning (MARIANI, 1956b)], to more or less regular drift, to the earth's curvature—especially at twilight, etc.

A difficulty arises in our model if we consider the geomagnetic control of the *F2*-layer. However, as we have already said, our "static" theory has to be corrected by taking account of tidal effects, which are subjected to this very geomagnetic control and may be remarkable even if a small drift velocity is assumed. On the other hand, the diffusion is also controlled by the geomagnetic field. Lastly, we certainly cannot exclude the possibility of some geomagnetic control of the ionospheric physical conditions, particularly of the height and latitudinal temperature distribution.

Finally, we note that our model could provide some useful suggestion for an explanation of the north–south asymmetries we have found (MARIANI, 1959) in the worldwide seasonal and non-seasonal variations of the electron density in the *F2*-layer. These asymmetries might be due to some slight difference in the thermal characteristics of the two hemispheres, which could cause a latitudinal asymmetry in some parameter of the theory (for example s or σ).

Future researches on the effective geometrical height distribution of electron density, on its diurnal, seasonal and latitudinal variations and on its drift movements will be very important for testing our hypotheses. Besides, further phenomenological and theoretical study of our model, in other physical conditions, could be useful.

Acknowledgements—I gratefully acknowledge the help I have received from the many individuals who have furnished me with ionospheric data.

REFERENCES

BARAL S. S. and MITRA A. P.	1950	*J. Atmosph. Terr. Phys.* **1,** 95.
BATES D. R.	1954	*Rocket Exploration of The Upper Atmosphere* pp. 347–355. Pergamon Press, London.
BATES D. R. and MASSEY H. S. W.	1946	*Proc. Roy. Soc.* A **187,** 261.
BRADBURY N. E.	1938	*Terr. Magn. Atmos. Elect.* **43,** 55.
BURKARD O.	1956	*J. Atmosph. Terr. Phys.* **8,** 83.
DOMINICI P. and MARIANI F.	1955	*Ann. Geofis.* **8,** 103.
DOMINICI P. and MARIANI F.	1956	*Nuovo Cim.* Suppl. **4,** 1589.
HAVENS R. J., FRIEDMANN H. and HULBURT E. O.	1954	*Report of Cambridge Conference* p. 237, Physical Society, London.
HIRONO M. and MAEDA H.	1954	*J. Geomagn. Geoelectr., Kyoto* **6,** 127.
HIRONO M. and MAEDA H.	1955	*Rep. Ionos. Res. Japan* **9,** 86.

VOL. 16 1959

Jackson J. E.	1956	*J. Geophys. Res.* **61,** 107.
Kallmann H. K., White W. B. and Newell H. E., Jr.	1956	*J. Geophys. Res.* **61,** 513.
Mariani F.	1955	*Ann. Geofis.* **8,** 59.
Mariani F.	1956a	*Ann. Geofis.* **9,** 43.
Mariani F.	1956b	*Ann. Geofis.* **9,** 219.
Mariani F.	1956c	*Ann. Geofis.* **9,** 245.
Mariani F.	1959	*Nuovo Cim.* In press.
Martyn D. F. and Pulley O. O.	1936	*Proc. Roy. Soc.* A **154,** 455.
Mitra A. P.	1951	*J. Atmosph. Terr. Phys.* **1,** 286.
Mitra A. P.	1952	*Indian J. Phys.* **26,** 79.
Ratcliffe J. A.	1956	*J. Atmosph. Terr. Phys.* **8,** 260.
Sato T.	1954	*J. Geomagn. Geoelectr. Kyoto,* **6,** 99.
Sato T. and Namikawa T.	1954	*J. Geomagn. Geoelectr. Kyoto,* **6,** 157.

OL.
16
959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 174 to 178. Pergamon Press Ltd. Printed in Northern Ireland

Molecular oxygen densities in the mesosphere at Fort Churchill

J. E. KUPPERIAN, Jr., E. T. BYRAM and H. FRIEDMAN
U.S. Naval Research Laboratory, Washington 25, D.C.

(*Received* 7 *October* 1958)

Abstract—The transmission of solar radiation to ultra-violet photodetectors in a rocket was used to obtain O_2 densities in the mesosphere. Densities between 70 and 86 km were lower by a factor of 1·8 in early spring than in midsummer. A temperature minimum of 186°K was observed at 83 to 85 km in March. Dissociation appeared to begin near 86 km in July and 96 km in March.

THE change in overhead mass of atmospheric gases can be determined by measuring the variation in solar radiation received by a rocket as it traverses the atmosphere. From such transmission data it is possible to deduce the altitude variation of molecular concentrations. At any altitude the optical depth of an atmospheric constituent is the product of its absorption cross-section and the number of molecules per square centimeter column above that altitude. If a solar wavelength is chosen which is absorbed relatively strongly by only one constituent, its concentration can be accurately determined for a range of plus or minus a scale height above and below the altitude of unit optical depth. In the ultra-violet between 1000 Å and 2000 Å, virtually all the absorption is attributable to molecular oxygen and this range of the spectrum has been utilized to determine the O_2 concentration (BYRAM *et al.*, 1955). At X-ray wavelengths near 50 Å all constituents, whether molecular or atomic, absorb independently of chemical combination and it is possible to determine the total air density without knowing composition (BYRAM *et al.*, 1956). The present paper is concerned with measurements of the O_2 density in the mesosphere at Fort Churchill in midsummer of 1957 and at the end of the 1957-58 winter. The data obtained were sufficiently accurate to demonstrate a seasonal variation in concentration and a change in the level at which dissociation begins.

In a typical experiment, the photo-ionization chamber is mounted on the skin of the rocket looking outward. The rocket is fired and the ion chamber current is telemetered continuously as the rocket climbs. Fig. 1 illustrates the kind of record usually obtained. When the rocket is deep in the earth's atmosphere there is no response because the overhead air absorbs the solar ultra-violet radiation completely. As the rocket approaches the upper reaches of the atmosphere the detector begins to see radiation and at unit optical depth the rate of change of transmission reaches a maximum. From there on, the intensity changes less rapidly as the rocket climbs above the absorbing atmosphere. The undulating responses are due to the roll of the rocket, since the detector produces signals only when it sees the sun. Amplitude variations, after the rocket is "above the atmosphere", are produced by precession of the rocket which changes the angle of incidence of solar radiation on the window of the ion chamber.

To determine the altitude variation of radiation absorption we must correct the individual responses to some constant aspect angle (angle between normal to

VOL. 16 1959

detector window and direct line to sun). The aspect correction is obtained throughout flight from the signals of visible light photocells having known angular responses to sunlight.

Between 1000 Å and 2000 Å three wavelength regions have been utilized for absorption measurements. With an overhead sun unit optical depth for 1500 Å radiation falls in the *E*-region; Lyman-α, which fits perfectly the deep narrow window at 1215–1217 Å, reaches unit optical depth in the *D*-region; and wavelengths near 1300 Å fill the altitude range midway between the *D*- and *E*-regions.

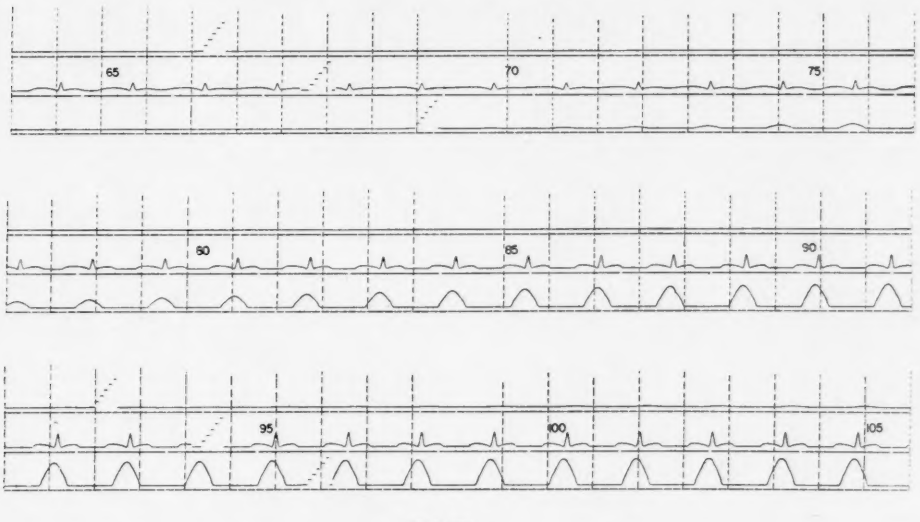


Fig. 1. Portion of telemetering record of IGY Aerobee 3·19F showing the signals as a function of altitude in kilometers above sea level. The upper channel shows the CaF_2 chamber response and the lower channel the LiF chamber response. Peaks in ion chamber currents occur when the chamber windows face the sun. The middle channel is the signal from the aspect photocells. The LiF detector is responding primarily to the Lyman-α flux while the CaF_2 detector responds only to the softer radiation above 1225 Å.

With the sun near the horizon the absorption levels all rise about 2 scale heights.

In principle it is possible to fly a grating monochromator and a photocell to measure the transmission of monochromatic radiation. In practice, a much simpler method is feasible, utilizing narrow-band photon counters or ionization chambers. The technique of preparing such detectors sensitive to the ranges of 1425–1500 Å, 1225–1350 Å and 1100–1350 Å has been described by CHUBB and FRIEDMAN (1955). The first of these fits almost flat maximum of the Schumann dissociation continuum where the absorption cross-section of O_2 is almost constant at about $1{\cdot}4 \times 10^{-17}$ cm². In the second range of wavelengths the cross-section is about 10^{-18} cm². At Lyman-α the cross-section reaches a minimum of about $0{\cdot}8 \times 10^{-20}$ cm². The latter two ranges are useful in determining the O_2 concentration of the mesosphere.

Fig. 2 shows the O_2 densities derived from two flights at Fort Churchill, Canada, 59°N using IGY rockets NN3. 13F and NN3.19F. The respective launching times were 1600 CST, 29 July 1957 and 1207 CST, 23 March 1958. The raw data for flight 3.19F consisted of intensity vs. altitude measurements spaced with the

roll period of the rocket 0·8 sec from two ion chambers covering 1065–1350 Å (lithium fluoride window) and 1225–1350 Å (calcium fluoride window). The data from the CaF_2 chamber were interpreted in terms of the solar spectral distribution photographed by JOHNSON *et al.* (1958) and the absorption cross-sections of WATANABE *et al.* (1953). The same data were also used to correct the response of the LiF chamber for radiation other than Lyman-α. This correction amounted to 3·3 per cent of the integrated intensity in the 1065–1350 Å band. An additional

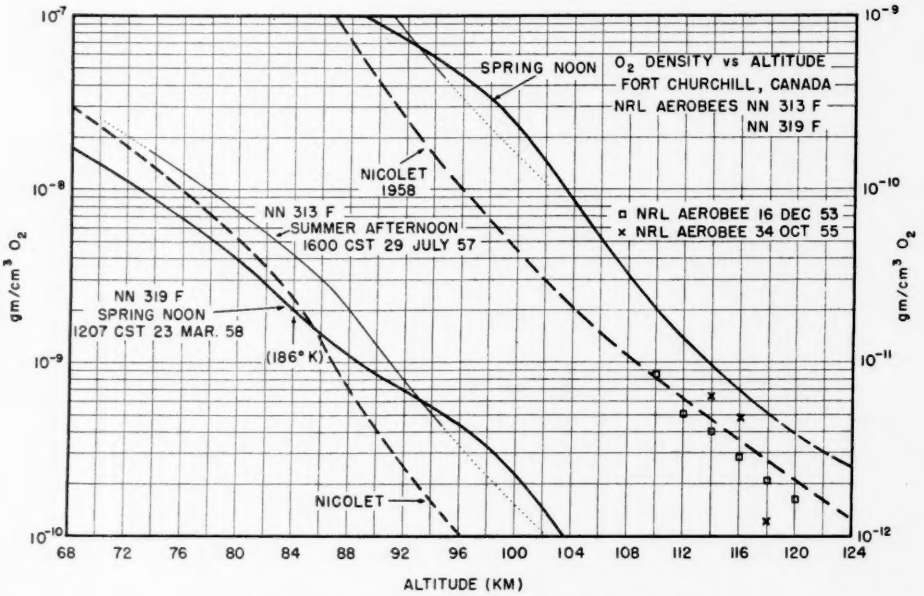


Fig. 2. O_2 density vs. altitude at Fort Churchill, Canada.

correction of 2·5 per cent was made to the "above the atmosphere" response of the LiF chamber to account for the Si III line at 1206·5 Å which has a high absorption cross-section. The contribution of this line appears in the data at altitudes above all measurable increase in the Lyman-α radiation and its variation at altitudes greater than 120 km is in accord with the absorption cross-section of $1{\cdot}7 \times 10^{-17}$ cm² published by WATANABE *et al.* (1953).

The corrected Lyman-α data were used to determine total vertical centimeters of O_2 at s.t.p. vs. altitude using an absorption coefficient of 0·22 cm^{-1} (s.t.p.) according to DITCHBURN *et al.* (1954) and a cosecant correction of 2·00 for solar elevation. The data from the CaF_2 chamber, which represent the average response of several weak emission lines of known absorption coefficients, were used in a similar manner to extend the analysis to 125 km. Since the 3·13F flight was not instrumented with the CaF_2 chamber, its data for the LiF chamber were adjusted with the same corrections derived from 3.19F for corresponding optical depths. As a result the O_2 analysis for 3.13 is only good for the altitudes from 70 to 100 km. The cosecant correction for 3·13F was 1·79.

VOL. 16 1959

The accuracy of each curve of intensity vs. altitude depends on a number of factors, such as noise and cross-talk in the telemetering record and uncertainty in the aspect correction. We estimate that the errors introduced by all sources in flight 3·13 leave an uncertainty of about ± 10 per cent in the final densities. The data from 3·19 is reliable to better than ± 5 per cent in the individual points and still better in the smoothed final curve.

The curves of intensity versus altitude were differentiated to produce Fig. 2. The curves are drawn as heavy lines where the scale height derived from the slope at a given altitude is quantitatively significant. The dotted portions are not extrapolations, but indicate less reliable data. NICOLET's theoretical curve (1958) is included for comparison. It corresponds very closely to ROCKET PANEL data (1952) below 75 km.

CONCLUSIONS

Several features of the O_2 distribution are clearly indicated by the curves of Fig. 2.

(1) The O_2 density on the two flights differed in the 70–86 km region by about a factor of 1·8, being highest in the summer. Since dissociation is negligible at these altitudes, the same factor applies to the density of the total atmosphere.

(2) The "dissociation level" appears as a rapid change in scale height beginning at about 86 km in the mid-summer flight and 96 km in the early spring flight. A difference of 10 km in the "dissociation level" can be understood in terms of the seasonal change in average solar elevation. This effect should be greater at Fort Churchill's latitude than at White Sands.

(3) The temperature minimum for the flight 3·19 in March is 186°K (+4, −8°) at an altitude of 83–85 km. It is unlikely that dissociation has an appreciable effect on the scale height from which this temperature is derived. (At all altitudes the scale height determined from the O_2 data must be lower than the atmospheric density scale height if any dissociation occurs.

(4) There appears to be an increase in scale height and therefore temperature (3.19) above the temperature minimum, before dissociation introduces a rapid decrease. This indication of temperature rise would follow the atmospheric gradient into *E*-region if not for dissociation.

(5) The high altitude distribution of O_2 from the Fort Churchill data is consistent with White Sands data (BYRAM *et al.*, 1957) from NRL Aerobees 16 and 34.

(6) The lower density at 70 km for the early flight implies a colder winter than summer atmosphere below this level. This is in agreement with WHIPPLE's interpretation of meteor data (WHIPPLE *et al.*, 1949).

(7) Below 80 km, the Rocket Panel White Sands density is within the spread of the two Fort Churchill measurements.

REFERENCES

BYRAM E. T., CHUBB T. A. and FRIEDMAN H.	1955	*Phys. Rev.* **98,** 1594.
BYRAM E. T., CHUBB T. A. and FRIEDMAN H.	1956	*J. Geophys. Res.* **61,** 251.
BYRAM E. T., CHUBB T. A. and FRIEDMAN H.	1957	*Threshold of Space* p. 211. Pergamon Press, London.
CHUBB T. A. and FRIEDMAN H.	1955	*Rev. Sci. Instrum.* **26,** 493.

DITCHBURN R. W., BRADLEY J. E. S., CANNON C. G. and MUNDAY G.	1954	*Rocket Exploration of the Upper Atmosphere* p. 327. Pergamon Press, London.
JOHNSON F. S., MALITSON H. H. PURCELL J. D. and TOUSEY R.	1958	*Astrophys. J.* **127,** 80.
NICOLET M.	1958	*Scientific Report* 102. Ionospheric Research Laboratory, Pennsylvania State Univ.
ROCKET PANEL	1952	*Phys. Rev.* **88,** 1027.
WATANABE K., INN C. Y. and ZELIKOFF M.	1953	*J. Chem. Phys.* **21,** 1026.
WHIPPLE F. L., JACCHIA L. and KOPAL Z.	1949	*The Atmosphere of the Earth and Planets* (Edited by KUIPER G. P.) p. 149. University of Chicago Press.

VOL.
16
1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 179 to 198. Pergamon Press Ltd. Printed in Northern Ireland

RESEARCH NOTES

Transient fine structure of the *E*-layer

(*Received* 25 *March* 1958)

In rocket experiments a rather complicated fine structure of the *E*-layer was observed by several groups (AFCRC 1953, NRL 1954, "Skylark" 1958), and it was maintained that it did not show up on the ground-based ionograms taken simultaneously. It appears to me that that discrepancy is not a fundamental fault of ground-based echo sounders but lack of sensitivity of the specific types used for comparison.

At Lindau the fine structure of the *E*-layer is being observed on routine ionograms since the start of an improved ionosonde in 1950. For example a typical ionogram taken in 1952 (Dieminger, 1954) shows five stratification levels between *E* and *F* (Fig. 1). The ionosonde used at Lindau has a peak power of 10 kW and sweeps the frequency range of from 1 to 16 Mc/s within 8, 4 or 2 min. The pulse repetition frequency is 50 c/s, the pulse duration 80 μsec. The antenna system consists of three vertical rhombics used for transmission and reception simultaneously. The normal mode of operation is with 4 min sweeps since a loss of details is experienced as soon as the frequency variation is speeded up to 2 min sweeps. It can be shown that that loss is not due to the detuning effect but to the deterioration of the integrated signal–noise ratio. That ratio, however, becomes worse still for panoramic ionosondes with a total sweep duration of 15 sec or less. Since the selective absorption is very pronounced in the transition region between *E* and *F* it is not surprising that on panoramic ionograms a gap shows up where an equipment with high power and slow frequency variation gives a multitude of stratifications.

The variability of the fine structure of the *E*-layer may be demonstrated by Fig. 2. The records were taken with our standard ionosonde sweeping the subrange between 2·8 and 4 Mc/s only in quick succession (30 sec). The fine structure of the echoes between f_0E and $f_{\min}$ *F2* is very complicated and varies from one record to the other. Apparently the variation of virtual height of the intermediate echoes is due to retardation which in turn varies considerably with the vertical gradient of electron density within the stratification. The multiple structure of the retarded *F2*-trace cannot be explained by vertical echoes alone. Doubtless off-vertical reflections are playing an important role.

The phenomena may be explained best by a region where the vertical gradient of electron density varies locally by small amounts, and which drifts over the point of observation. That picture implies of course horizontal gradients of electron density and tilts of the individual iso-ionic planes.

The records are taken at random, and are typical for normal conditions in winter-time. The degree of stratification, however, is subjected to large variations. A thorough investigation of the phenomena is in progress. Full details will be published in due course.

W. Dieminger

Max-Planck-Institut für Aeronomie
Institut für Ionosphären-Physik
(20b) *Lindau über Northeim/Han., Germany*

References

Dieminger W. 1954 *Physics of the Ionosphere* p. 53. Physical Society, London.

Solar activity and the altitude of the tropopause near the equator

(*Received* 6 *April* 1959)

THE study of the behaviour of the tropopause at Leopoldville ($\varphi = 4{\cdot}4°S$, $\lambda = 15{\cdot}0°E$) during the period of radio soundings from June 1953 to December 1958 reveals a very good correlation between its pressure level or altitude and the Zürich sunspot number R of solar

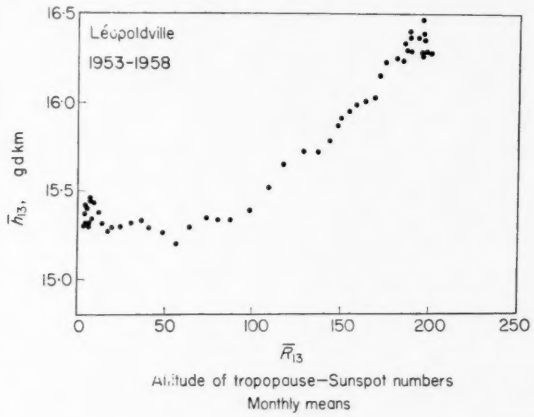


Fig. 1.

activity. It can be seen from Fig. 1 that the running means over 13 months of the height of the tropopause, being at 15·3 km in December 1953, begin to rise at the end of 1955 up to 16·4 km, reached at the end of 1957. This increase coincides with an increase of the running

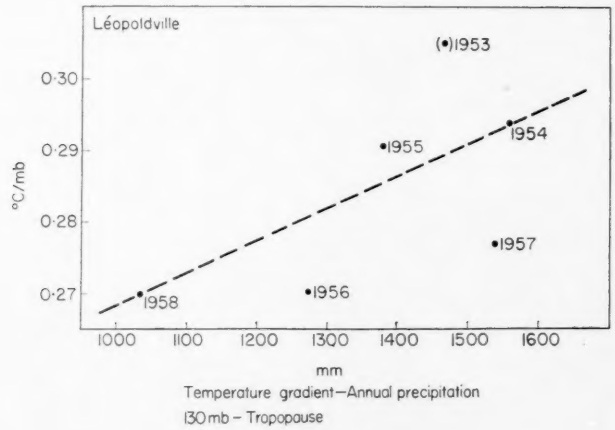


Fig. 2.

means over 13 months of sunspot number from $R_{13} = 70$ to $R_{13} = 200$ by the end of 1957. No regard has yet been given to a possible phase lag between the two phenomena. An investigation of the temperature at the tropopause level shows its constancy throughout the 5 years from sunspot minimum to maximum, so that the average temperature of —79·8°C for the 5 years moves up from 15·2 to 16·4 km height. It is obvious that this displacement

VOL. 16 1959

TOL.
16
959

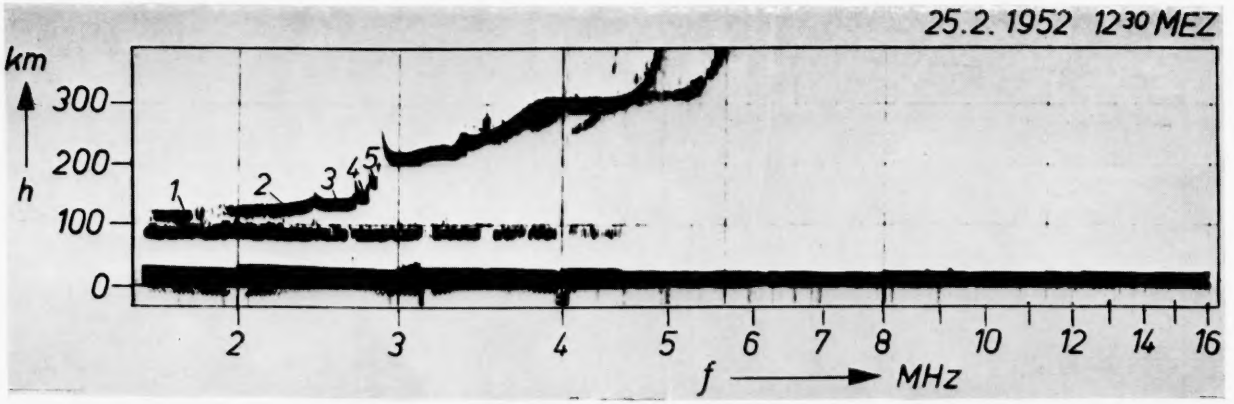


Fig. 1. Lindau ionogram with expanded time base from 1, 2, 3, 4, 5 stratifications in the *E*-layer. (*See* DIEMINGER, p. 179.)

Fig. 2a.

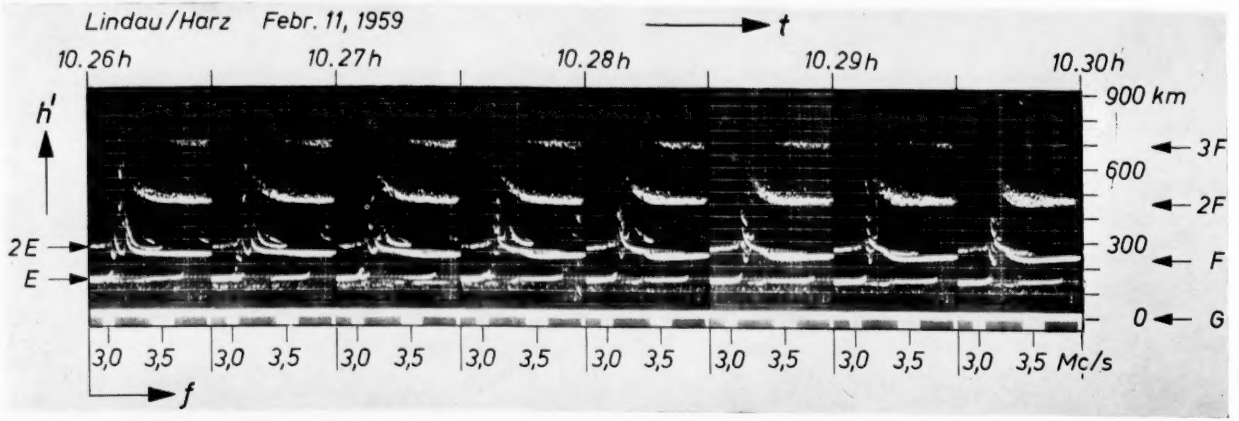


Fig. 2b.

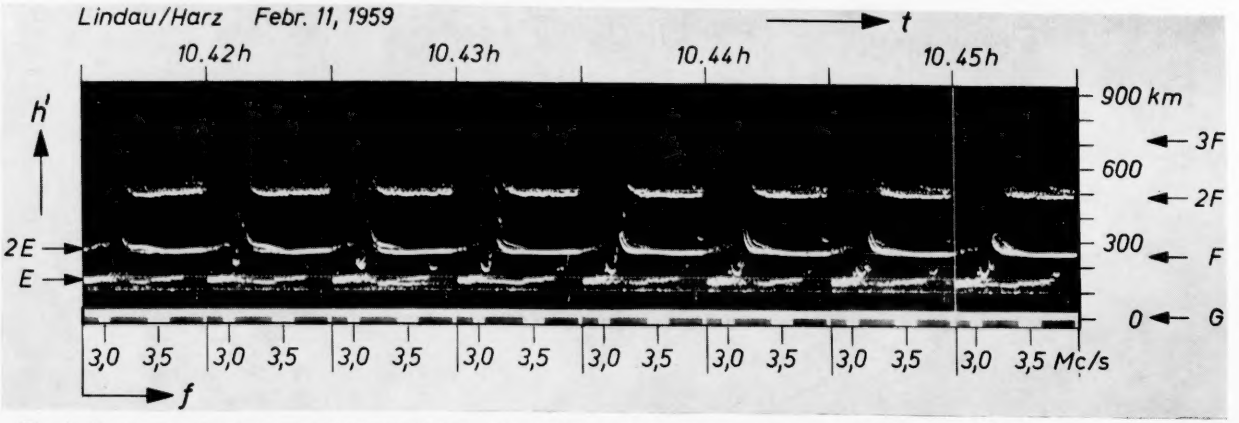


Fig. 2. Fast succession ionograms 2·8–4·0 Mc/s demonstrating rapid variations of E-layer stratification. (*See* DIEMINGER, p. 179.)

VOL. 16 1959

of the important discontinuity between the troposphere and the stratosphere, in close relation with the solar activity, will influence the vertical distribution of atmospheric temperature and, as a consequence, the weather phenomena too. We feel this is the first time that such a close correlation between sunspot number and the behaviour of the low atmosphere has been shown. Hitherto the influence of solar activity on tropospheric phenomena was supposed to be doubtful or negligible, and on the stratosphere or ozonosphere it was assumed to exist, but was not yet clearly shown. The change of the temperature

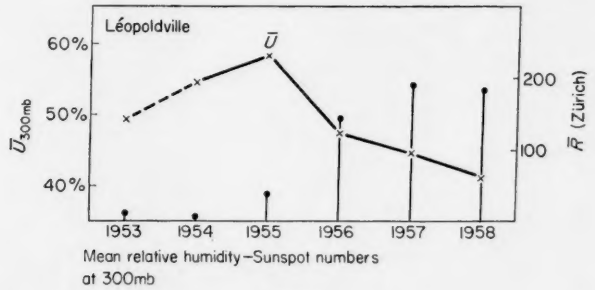


Fig. 3.

gradient in the uppermost troposphere from 1954 to 1958 reveals interesting effects on weather, one of which will be dealt with here. The mean vertical temperature gradients between 150 mb level and the tropopause for each year from 1953 (only June to December 1953 are available) and 1954 to 1958 are shown in Fig. 2 together with the total precipitation of the corresponding years at Leopoldville airport. The use of an average of rainfall upon a greater area would be more reasonable and will be tried later. The total precipitation is, in general, lower when the temperature decreases less with height, and this is the case in

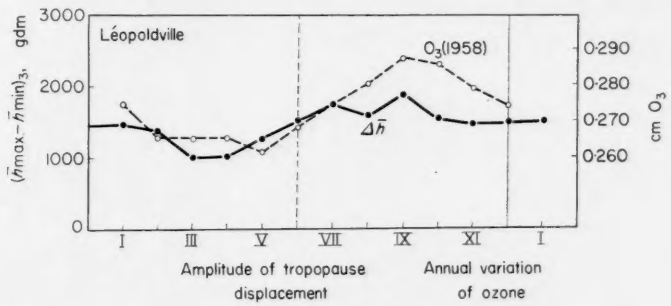


Fig. 4.

the years of high sunspot number due to the elevated tropopause. There is still a high random effect in the precipitation of a single station. The relative humidity, however, at the highest level regularly measured, i.e. at 300 mb, shows still much better correlation (Fig. 3). Water vapour in the atmosphere is transported upward from the ground, the amount being the greater, the greater the temperature gradient. This happens at sunspot minimum (1954–1955), while at sunspot maximum (1957–1958) the mean relative humidity is considerably reduced. It remains to be studied if this phenomenon is due to an actual deficit of water vapour or is a consequence of increased temperature at 300 mb.

Finally a plot of a running means over 3 months of the differences for each month, between the mean maximum height of tropopause at sunspot maximum and the mean

minimum height at sunspot minimum shows a definitely greater amplitude of displacement of this surface during the second half of the year. For this period the amplitude is 1800 gdm (or 12%) against 1000 gdm (or 7%) for the first half of the year. This change in amplitude due to the solar activity upon the altitude of the tropopause is in fairly good agreement with the yearly variation of total amount of ozone above Leopoldville (available for 1958 and assumed similar for the other years). Its highest value appears in September, too, and its lowest in April/May. This result (Fig. 4) may give a hint towards a probable mechanism involving the solar activity upon the ozone layer with further transfer downward to the tropopause. A detailed investigation of other phenomena will be published later elsewhere.

Section de Géophysique
Service Météorologique
Léopoldville

D. STRANZ

Ionospheric observations on the *F*-region during the solar eclipse of 19 April 1958

(*Received* 31 March 1959)

VOL. 16 1959

THE general features of the solar eclipse effects in the *E*- and *F1*-layers are now well known. For the *F2*-layer, however, it is not so. It is, in fact, even doubtful if any simple and regular eclipse effect (such as a decrease in the critical frequency) exists for the *F2*-layer. Therefore observations of the eclipse effect in the *F2*-layer, whenever opportunity occurs, are useful. Such an opportunity presented itself to us on 19 April 1958 at Calcutta when there was a solar eclipse with 61 per cent obscuration. Observations were made during this eclipse at the Ionosphere Field Station, Haringhata (lat. 22·9° N, geomag. lat. 12·5° N) with the routine sweep-frequency equipment, an Australian type automatic recorder, having a frequency sweep from 1·0 to 13·0 Mc/s covered in 1 min 55 sec and peak power of about 1 kW.

Observations on the eclipse day were taken every 2 min, whereas those on the control days were at intervals of 10 min. Accurate measurement could not be made at frequencies less than about 4·0 Mc/s on eclipse day due to interference. This restricted the eclipse observations to the *F*-layers only.

F2-*layer*. Fig. 1 shows the variation of maximum electron density (N_mF2) of the *F2*-region on the eclipse day (full line). For comparison, the mean variation of the same, as obtained from ten control days centred on the eclipse day, is also given (broken line). It will be seen that although at the beginning of the eclipse, N_mF2 of the layer was greater than its average control day value, it decreased during eclipse and reached a minimum about 20 min after the eclipse maximum. The minimum N_mF2 was about 26 per cent of the average control day value at the same hour. However, on account of the great variability in the behaviour of the *F2*-layer even on magnetically quiet days, one may hesitate to ascribe this decrease to simple eclipse effect specially because of the low value (about 61 per cent) of the maximum obscuration of the sun's disk. Hence to test further if the "dip" was an eclipse effect we calculated the total electron content (n) in a unit column in the *F*-layer both during eclipse as also on the control days by numerical integration from the true height–electron density distribution curve. The true height analysis of the eclipse h'–f records was made by the method of KELSO (1952) neglecting earth's magnetic field.

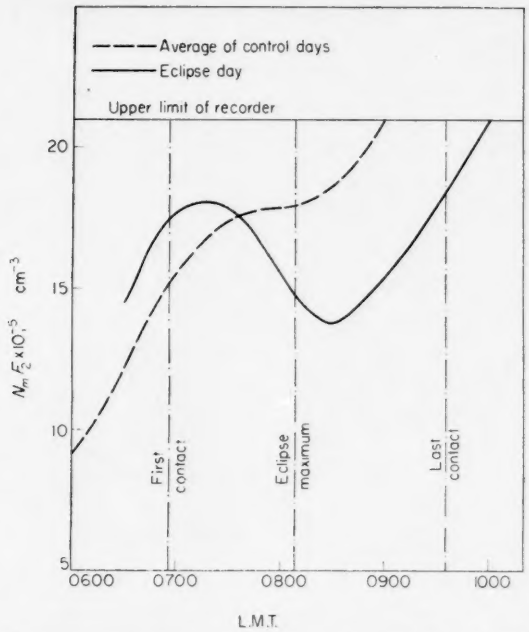


Fig. 1. Variation of maximum electron density (N_mF2) in the $F2$-region. The line at $N_mF2 \times 10^{-5} = 21$ corresponds to the upper limit of the recorder.

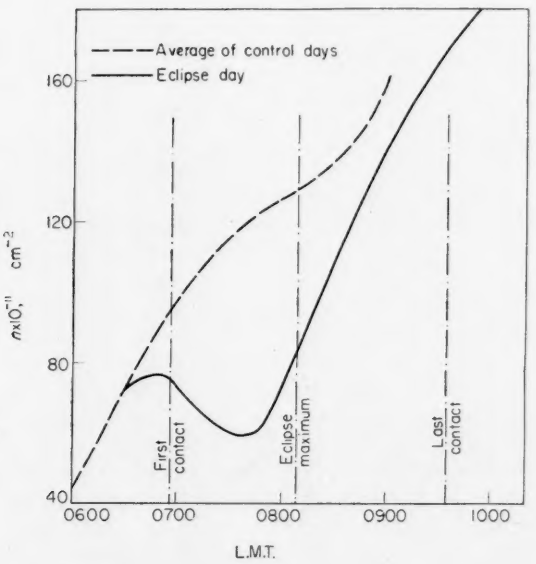


Fig. 2. Variation of "total electron content" (n) in a unit column of the $F2$-layer.

The result is shown in Fig. 2 in which the n–t curve for the eclipse day (full line) is compared with the corresponding average control day curve (broken line). It will be seen that the dip is present in the former curve only, and not in the latter. One should also consider in this connexion the fact that the eclipse took place during the morning hours when the ionization density in the layer was building up. During this period of the day N_mF2 varies more or less in regular manner, although in certain other parts of the day the changes may be random. As a matter of fact, the control day values during the morning hours did not vary greatly from day to day and such a prominent dip at this hour was not observed on any of the control days.

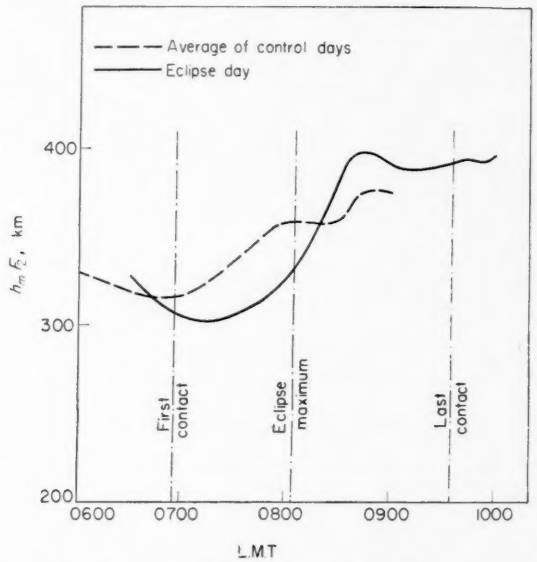


Fig. 3. Variation of the height of the $F2$-layer peak (h_mF2).

Another important parameter to be studied during an eclipse is the true height of the $F2$-layer peak (h_mF2). Its variation during the eclipse under study is given in Fig. 3. It will be noted that h_mF2 remained above 300 km all the time. It might be mentioned, in this connexion that RATCLIFFE (1955) has made an estimate of the time constant of the life of the electrons at different heights in the $F2$-layer and has shown that an eclipse of average duration of about 2 hr is not likely to have any appreciable effect on the electron density in the layer at heights above 300 km approximately. Our results, however, do not agree with this, since the electron density in the layer was affected.

F1-*layer*. There were some indications of the appearance of an $F1$-peak; but the peak did not appear sufficiently prominently to allow it to be scaled as such. This is partly because the eclipse took place in the morning at an hour at which bifurcation of the F-layer is not expected. Further, in the present epoch of the sunspot cycle (i.e. near the maximum of the cycle) clear bifurcation of the F-layer is generally not observable. We, however, examined the incomplete $F1$-peak and the part of the trace below it in the sequence of ionograms. It is known that $F1$-peak sometimes becomes more prominent(RATCLIFFE, 1956) during the solar eclipse than on control days and the height of the $F1$-layer increases considerably at times. On the present occasion, however, no such effects were observable; the height of the $F1$-layer ($h'F$) remained nearly constant at about 250 km.

VOL. 16 1959

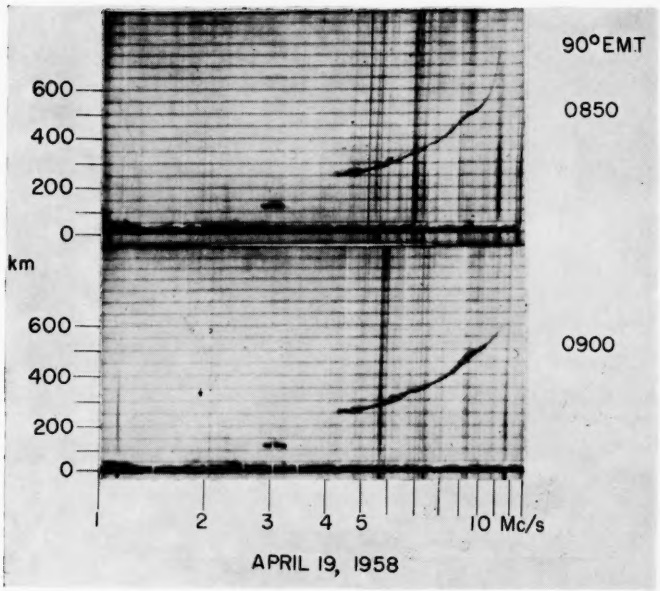


Fig. 4. Nature of the $F1\frac{1}{2}$-peak observed near the end of the eclipse.
DATTA *et al.*, p. 184

VOL.
16
1959

F1½-*layer*. Attention has recently been drawn (RATCLIFFE, 1955) to the fact that during an eclipse an additional peak may appear in *F*-layer between the *F1*- and *F2*-peaks. This is described as an *F1*½-layer. The peak at about 9·4 Mc/s in the ionogram (Fig. 4) might appear to be such a peak during the eclipse under study. It is, however, doubtful if this peak is an eclipse effect, since similar peaks were observed at about the same hour on control days as well.

Acknowledgements—The work forms part of the programme of the Radio Research Committee of the Council of Scientific and Industrial Research. The writers are grateful to Professor S. K. MITRA for his encouragement and advice. They are also thankful to Dr. A. K. SAHA under whose supervision the work was carried out.

S. DATTA
P. BANDYOPADHYAY
R. N. DATTA

Institute of Radiophysics and Electronics
University of Calcutta
Calcutta, India

REFERENCES

KELSO J. M. 1952 *J. Geophys. Res.* **57**, 357.
RATCLIFFE J. A. 1955 *Solar Eclipses and the Ionosphere; Survey paper.* Pergamon Press, London.
RATCLIFFE J. A. 1956 *J. Atmosph. Terr. Phys.* **8**, 260.

Transient modes of high-frequency radio wave propagation across the auroral zone

(*Received* 17 *April* 1959)

SWEEP frequency oblique incidence radio soundings of the ionosphere across the auroral zone have been made regularly since October 1958. The path is between Winnipeg (49·9 N, 97·4 W) and Resolute Bay (74·7 N, 94·9 W), a distance of 2760 km, and the mid-point is located in the zone of maximum visual auroral occurrence.

The records obtained at both ends of the circuit indicate extra reflections which produce one or more loci that have the same general structure as fragments of the locus produced by regular ionospheric reflections. These traces are much more transient in nature than the traces for regular one-hop reflection. The maximum frequencies propagated on these extra modes may be either greater or smaller than the *F*-layer m.u.f., but they are usually greater. These extra traces are present on from 50 to 60 per cent of the recordings for the months of November and December 1958.

Fig. 1 is an example showing two extra traces with both low and high angle rays visible. Fig. 2 and 3 show examples of an extra high angle ray trace and low angle ray trace, respectively. The type of trace shown in Fig. 3 that crosses the one-hop high angle ray trace and extends to a higher frequency than the one-hop m.u.f. has also been observed on a path from Sodankylä to Lindau (MÖLLER, 1958).

The occurrence of these extra traces may be due to irregularities in the auroral zone ionosphere that act as scattering sources. An irregularity situated between the transmitter and the receiver can provide two possible types of deviate paths: one in which the ray is reflected from the *F*-layer to the irregularity, from which the energy is scattered to the

receiver; another in which the ray is transmitted directly to the irregularity, and the energy is scattered forward into an F-layer reflection before arriving at the receiver. For each of these deviate paths, there may be a high and low angle ray path. A third possibility consists of those ray paths by which propagation occurs directly to the irregularity and then directly to the receiver. These three possibilities are adequate to explain the change in the upper frequency limits and the increase of the time delays over those of the regular one-hop mode. These extra traces can also occur as a result of either large-scale irregularities or folds in the ionosphere. However, these other models will not be discussed.

The characteristics of these traces will depend on the size and electron density of the irregularity as well as its position (horizontal range, lateral deviation, height). When the F-layer supports high angle and low angle reflections for both deviate paths, two extra traces would appear on the record as illustrated in Fig. 1. If the irregularity is located in the vicinity of the transmitter or the receiver, the density of the F-layer may support reflections on only one of these deviate paths. When the irregularity is located in the F-layer, direct ray reflections on the deviate path may be excluded by reflections at the bottom of the F-layer and only the high angle ray trace would appear, as illustrated in Fig. 2. The effect of a lateral deviation of the irregularity would be mainly to increase the time delays and the upper frequency limits of the traces. If the irregularity is located below the F-layer and far from the vertical through the mid-point of the circuit, propagation by means of the irregularity alone should occur and produce a trace such as the one shown in Fig. 3.

An attempt has been made to determine the position of the irregularity by using the traces on the records. In order to verify the proposed model some method must be found to resolve the ambiguities of the horizontal range and the height of the irregularity.

The oblique incidence sounders on the Winnipeg–Resolute Bay circuit are operated by the Department of Transport for the Defence Research Board.

B. J. Fulton
L. E. Petrie
W. S. P. Ward

Defence Research Telecommunications Establishment
Defence Research Board
Shirley Bay, Ottawa, Canada

VOL. 16 1959

Reference

Möller H. G. 1958 *J. Atmosph. Terr. Phys.* **13**, 173.

The ionospheric *E*-layer at Cape Hallett

Because of its latitude (72°18′S) Cape Hallett is well placed for studying the variations with solar zenith angle of f^0E between the range of 1–2 Mc/s.

In this frequency range it is fairly easy to separate normal ionization caused by ultraviolet light, and the contributions from particle influx, for they occur at different heights.

For this study, the ionograms from 20 April 1958 to 14 August 1958 were read very carefully, and only those E-layers unaffected by particle ionization were scaled. The values obtained were compared with those expected from a theoretical model of the atmosphere. The model was based on the following assumptions:

(1) f^0E at zero solar zenith angle is 4·0 Mc.

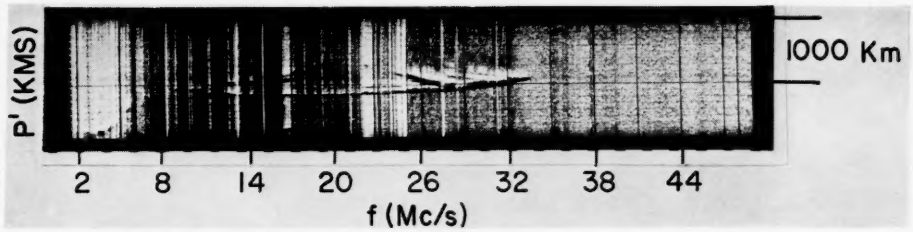


Fig. 1. Winnipeg–Resolute Bay; 30 October 1958, 1440 CST.

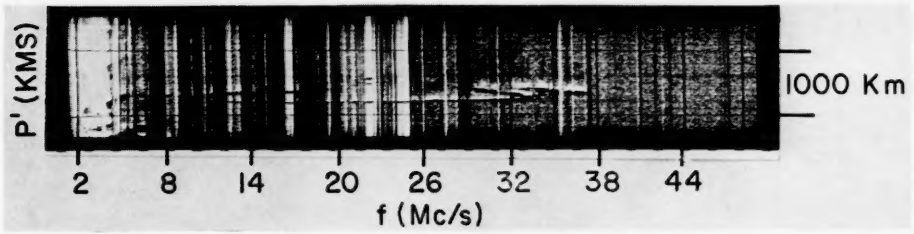


Fig. 2. Winnipeg–Resolute Bay; 19 November 1958, 1220 CST.

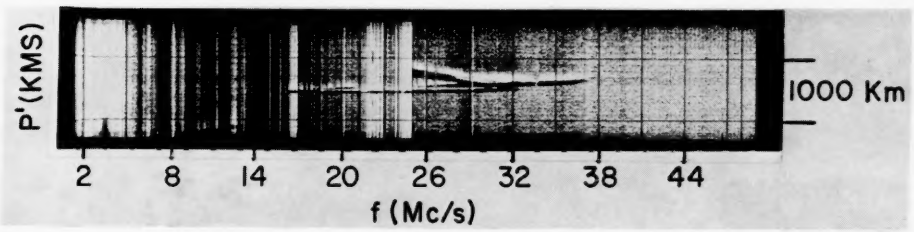


Fig. 3. Winnipeg–Resolute Bay; 2 November 1958, 1320 CST.

VOL.
16
1959

(2) There is equilibrium between ion formation and removal.

(3) The atmosphere has a scale height of 12 km at 100 km and 40 km at 240 km, with a linear gradient between.

(4) The appropriate Chapman function at each height is obtained by successive approximations (necessary because of the gradient of scale height).

(5) The coefficient of electron recombination is the same at all heights.

Table 1 shows the distribution of f^0E at the various expected values.

Table 1.

Expected value	1·0	1·1	1·2	1·3	1·4	1·5	1·6	1·7	1·8	1·9	2·0	2·1	2·2	Total
2·0										1	4	2		7
1·9									2	1	1	1		5
1·8									3	1	3			7
1·7							2	3	8	3	2			18
1·6							3	3	2	1				9
1·5					4	7	6	4	3					24
1·4			1	1	11	7	5	7	2					34
1·3	1	3	2	7	3	2	1							19
1·2	1	0	1	2										4

Observed value

Note. Fewer observations are recorded at times when f^0E is expected to be 1·6 Mc/s because of a regular diurnal appearance of particle activity.

The agreement between observed and expected values is good. The increased spread at the low frequency end seems to be due to interaction with the F-region.

A similar table based on a constant coefficient of attachment showed no agreement with the observed f^0E. This should not be construed as a verification of the recombination process, for a gradient of attachment coefficient could conceivably give a satisfactory behaviour.

Likewise, this study does not allow any conclusions about the atmospheric model adopted, as f^0E is not very sensitive to changes in it.

On the other hand $h'E$ gives a good test of the atmospheric model, but unfortunately the experimental data are not adequate.

G. A. M. King

Geophysical Observatory
Christchurch, New Zealand

The effect of the earth's magnetic field on m.u.f. calculations

(*Received 2 May* 1959)

Practically all of the methods at present used in the determination of maximum usable frequencies for ionospheric propagation neglect the effects of the earth's magnetic field.

THOMAS and ROBBINS (1958) have pointed out, however, that the inclusion of the magnetic field can appreciably affect the determinations of the electron density distribution and therefore the calculated m.u.f. It is the purpose of this note to draw attention to the influence of the earth's magnetic field on m.u.f. calculations using the transmission slider technique.

The Smith transmission curves (Circular 462) provide one of the most convenient and widely used methods for determining maximum usable frequencies. Direct measurements of m.u.f. (e.g. WIEDER, 1955) have shown that the observed values are, on the average, about 3 per cent higher than those obtained by the use of the transmission curves.

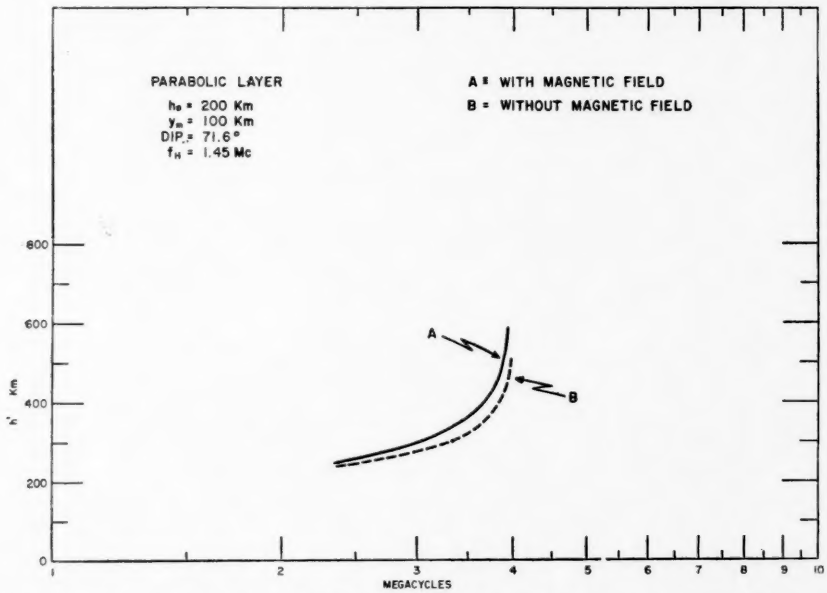


Fig. 1. The variation of virtual height of reflection with frequency in a parabolic layer.

If we consider any given electron density profile the equivalent heights calculated with the earth's magnetic field included are greater than those obtained when the magnetic field is neglected, except at the magnetic equator where they are equal for the ordinary ray. This is shown in Fig. 1 in which we see the equivalent height vs. frequency curves with and without the earth's magnetic field at middle latitudes in the case of a parabolic layer for which the base height $h_0 = 200$ km, half width $y_m = 100$ km, dip $= 71{\cdot}6$, and $f_c = 3{\cdot}97$ Mc/s. Now curve A represents the $h'f$ sweep which would be observed by a conventional ionosonde whereas curve B is the one appropriate for use with the Smith transmission curves. It is clear that application of the transmission curves to curve B will always yield higher m.u.f.s than those obtained from the observed $h'f$ curve. This is seen in Table 1 in which are shown the m.u.f.s obtained from curves A and B, respectively, for transmission distances of 1000 km, 2000 km and 3000 km. It is seen from Table 1 that the use of the appropriate $h'f$ curve increases the calculated m.u.f. by about 7·5 per cent.

This percentage difference is somewhat higher than the observed value. This may be due to the effect of the earth's magnetic field in decreasing the m.u.f. of an obliquely reflected ray. This latter point has been shown by the work of HASELGROVE (1957); however, it should be borne in mind that in the case of the observed m.u.f.s mentioned above,

VOL. 16 1959

the values of $Y(=f_H/f, f_H =$ the gyrofrequency) are of the order of 0·1 and HASELGROVE'S curves show that under these conditions the effect of the earth's magnetic field is small.

It would appear therefore that errors in m.u.f. calculations arise more by neglecting the effect of the magnetic field on the vertically incident ray rather than on the obliquely incident one. M.U.F. calculations may be improved by either of the two following procedures:

(a) Derive an electron density profile from each ionogram, then find the $h'f$ curve neglecting the magnetic field and apply the appropriate transmission curve.

Table 1

Distance (km)	1000		2000		3000	
Curve	*A*	*B*	*A*	*B*	*A*	*B*
M.U.F. (Mc/s)	5·85	6·30	9·30	10·0	11·50	12·35
Difference (Mc/s)	+0·45		+0·70		+0·85	
% Difference	+7·6		+7·5		+7·4	

A = With the magnetic field.
B = Without the magnetic field.

(b) Construct a new set of transmission curves, taking into account the earth's magnetic field.

Although the latter approach seems, at first sight, to be the simpler, it may require different sets of transmission curves for different regions of the world.

I am grateful to Mr. J. W. WRIGHT for furnishing the data on equivalent heights with the inclusion of the magnetic field.

K. DAVIES

National Bureau of Standards
Boulder, Colorado

REFERENCES

NATIONAL BUREAU OF STANDARDS	1948	*Ionospheric Radio Propagation*. Circular 462, Ch. 6.
HAZELGROVE J.	1957	*Proc. Phys. Soc.* B **70,** 653.
THOMAS J. O. and ROBBINS A.	1958	*J. Atmosph. Terr. Phys.* **12,** 77.
WIEDER B.	1955	*J. Geophys. Res.* **60,** 395.

Annual distribution of sporadic-*E*

(*Received* 4 *May* 1959)

THE geographical occurrence of sporadic-*E* ionization over North America has been studied utilizing the 19,000 or more reports of v.h.f. radio contacts. The reports were obtained during the calendar year 1949 and the results apply only to this period. Data were obtained mainly at a frequency of 50 Mc/s, which for the distances involved, corresponds to an equivalent vertical incidence frequency of about 8·5 Mc/s or higher. The reporting techniques and some analyses have been reported elsewhere.

From each radio contract, the location of the E_s cloud effective in allowing the radio wave reflection was determined. An annual summary was then prepared of the number of hours during which reflections occurred within each degree square. (A degree square is

defined as that area bounded by 1° of latitude and 1° of longitude.) The total number of hours that E_s was observed in given degree squares is summarized in Fig. 1. In this figure, the outermost curve embraces the entire area in which E_s was received. The numbered isopleths represent the occurrence of E_s in tenths of a per cent; i.e. E_s observed hours $\times$ 1000/8760.

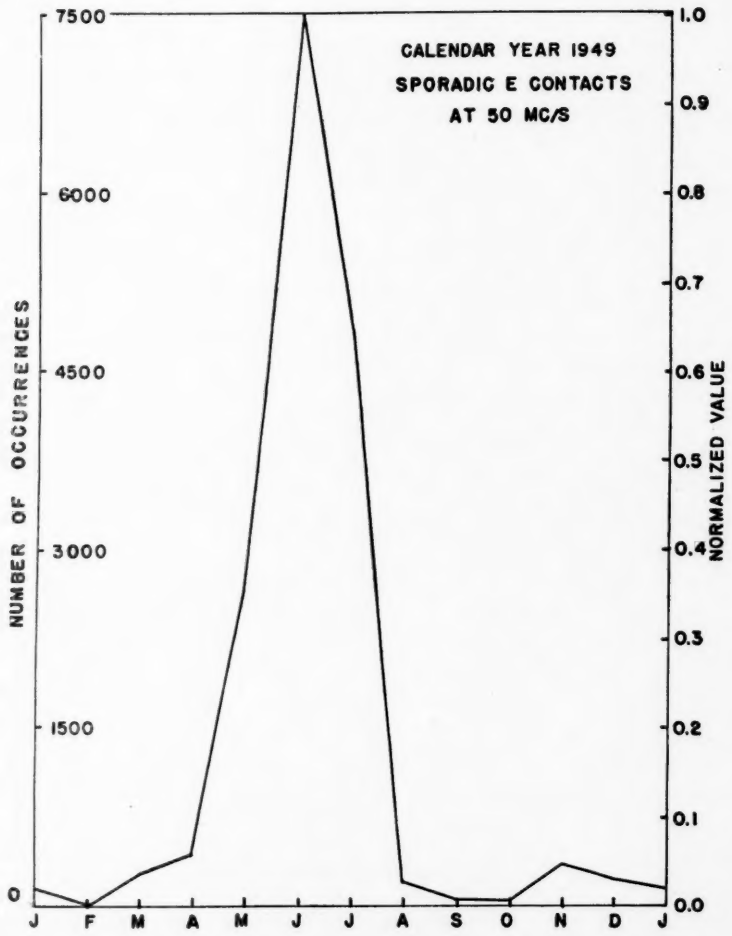


Fig. 2.

The diagram shows that practically every degree square within the observing area contained at least one occurrence of E_s during the calendar year. In the corners and along the periphery of the observing area, little or no E_s was reported. This apparent lack of reports indicates an inherent deficiency of the observational scheme; it does not lead to the conclusion that sporadic E was absent from these peripheral zones.

The reporting network allowed the detection of E_s within the outer curve. Within this entire region, the E_s occurrence attained a value of 0·5 per cent in two general areas: (a) south-western U.S., from northern California to Texas, and (b) central U.S., from

VOL. 16 1959

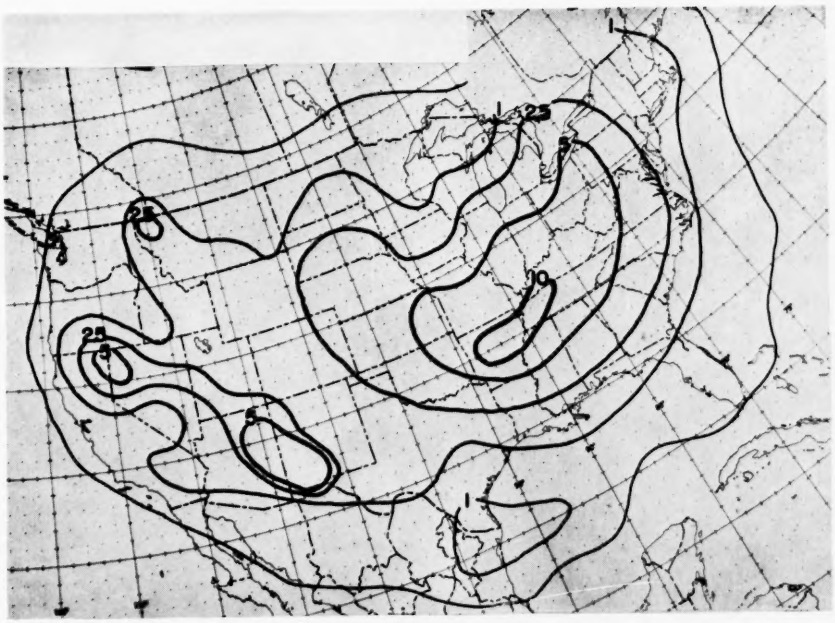


Fig. 1

VOL.
16
1959

New York to Oklahoma. The greatest occurrence, 1 per cent, was found in a relatively small band from Kentucky into Arkansas.

Some remarks may be made relative to the representativeness of the reporting network with respect to both time and geography. Fig. 2, portraying the monthly distribution of sporadic-E, reveals that the E_s distribution is markedly seasonal. The figure emphasizes that sporadic-E clouds capable of reflecting critical frequencies of about 8·5 Mc/s and higher is almost entirely a summer phenomenon in the northern hemisphere. Because E_s is essentially concentrated in the summer months, Fig. 1 provides a rough guide of the summer sporadic-E distribution. The effects of the remaining months introduces some discrepancies; however, these are considered minor.

The geographical bias introduced by the fixed location of the reporting network probably is confined mainly to small distortions. Such discrepancies undoubtedly stem from the uneven concentration of observers within the reporting areas. For example, the location of about two-thirds of the stations east of 95°W may account for the presence of the observed maximum E_s concentration in eastern U.S. (see Fig. 1). Some of the bias may be partially removed by assuming that E_s was distributed fairly uniformly over the observing area. On this basis it would be expected that E_s occurred about 1 per cent of the time (during 1949) within the entire observing area depicted in Fig. 1.

In extrapolating the results to other geographical regions two factors must be remembered. First, the sporadic-E observed in this study is predominantly middle latitude E_s, being neither the tropical nor polar variety. Secondly, from a large scale viewpoint, it seems reasonable to assume that within zonal circles around the globe, no marked (middle latitude type) sporadic-E gradients occur. It is accepted that some minor localizations may take place, (similar to the Aleutian Low, Bermuda High, etc., of tropospheric meteorology), but marked departures from the average would not be anticipated. Thus, to a first approximation, middle latitude E_s may be considered as being uniformly distributed in latitudinal zones around the earth. However, whether the zones are sensitive to specific geographic or geomagnetic influences still requires clarification.

On the basis of the preceding discussion, it is concluded that during 1949 the distribution of sporadic-E attained was 0·5 — 1 per cent over middle North America. Also, this concentration was probably found throughout the latitudinal zone 28°N — 55°N, with a possibility of some "islands" having the higher percentages.

N. C. Gerson

Trapels Road
South Lincoln, Mass.

Horizontal movements of visual auroral features

A number of isolated measurements of movements of individual features in an auroral display have been reported in the past, for example, Meinel and Schulte (1953). Recently Kim and Currie (1958) have given much more extensive results of a search for systematic movements, using time lapse photographs taken by all-sky cameras at three locations in Canada.

As a regular behaviour of radio echoes associated with aurora has been reported by many workers from all over the world, it is surprising that simple investigations of comparable visual movements have not been atempted earlier. All-sky cameras have been widely employed during the IGY, and there must now be a considerable amount of suitable information available. The material used here was obtained at Royal Society Base, Halley Bay, Antarctica (75°31′S, 26°36′W) during the southern winter 1956.

Changes in an auroral display which are appreciated visually will generally be those associated with changes of form, and they are consequently most impressive and most rapid during the short active period at the climax of a display. We wish now to concern ourselves with the drift of luminous clouds once formed, rather than with a changing pattern of excitation. We therefore adopt the criterion of KIM and CURRIE, which requires that the same readily identifiable feature shall be recognizable on at least four successive frames of the time lapse film. Frames were exposed once per minute during the period under investigation, but more recent films from Halley Bay have been exposed at four frames per minute and much more reliable information for this kind of analysis is available.

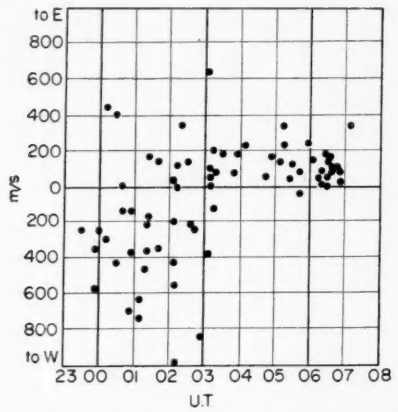


Fig. 1. Individual measurements of component of velocity along the geomagnetic parallels at Halley Bay during 1956 July to August. Ordinates: velocity in m/sec. Abscissae: universal time.

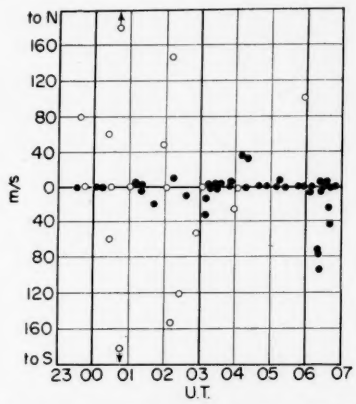


Fig. 2. Components of velocity parallel to the geomagnetic meridian through Halley Bay. Circles, ○, measurements on arcs. Spots, ●, other forms, mainly *DS*. Ordinates: velocity in m/sec. Abscissae: universal time.

A fixed height of 100 km over the earth is assumed in calculating the position of any feature. A grid is drawn with one set of lines spaced 10 km apart parallel to the geomagnetic parallels of latitude, and the other similarly spaced parallel to the geomagnetic meridian through Halley Bay. It is transformed first into azimuth and elevation from the point of observation, and then into the projection formed by the image in the all-sky camera. This grid is used on the screen of a microfilm reader to give linear north–south and east–west co-ordinates of features of interest on successive frames.

Arcs lying along the geomagnetic parallels permit of measurements of north–south movement only. Isolated diffuse surfaces permit measurement of movement in any direction and are much the most reliable features for this purpose. Quite faint areas, which would attract no attention visually and which occur during the quiet periods of the display, are the most useful. Rays are generally too short lived. They have been used on a small fraction of occasions, care being taken to refer the co-ordinates to the bottom of the ray which is likely to be near the assumed height. Locally bright areas, or folds, which move along otherwise quiet arcs may represent a different kind of motion. For the moment no distinction has been made between the results obtained from different features but it is obvious that the relative number of north–south and east–west movements is likely to be governed by the features which predominate at any given time.

The first graph shows individual east–west velocities against U.T. Local midnight is 0146 U.T. At Halley Bay 0302 U.T is the time when the azimuth of the sun at the equinoxes

VOL. 16 1959

is the same as that of the geomagnetic axis point. Systematic behaviour, strongly reminiscent of that observed by radar techniques, is revealed.

The second graph, for north–south movements, distinguishes between measurements on arcs and on other forms. For other forms the velocities measured are an order of magnitude less than the east–west velocities. Components of north–south geomagnetic motion considerably greater than those measured would be involved if the predominant motion were along the geographic parallels, or the local isoclinic lines, rather than along the geomagnetic parallels.

The analysis of the Halley Bay IGY films taken in the 1956 and 1957 winter seasons will extend the results to a wider range of hours, and resolve details of the behaviour suggested by these graphs.

S. Evans

Scott Polar Research Institute
Cambridge

References

Kim J. S. and Currie B. W. 1958 *Canad. J. Phys.* **36**, 160.
Meinel A. B. and Schulte D. H. 1953 *Astrophys. J.* **117**, 454.

Single and double inflexions on the *F*-trace

(*Received* 20 *May* 1959)

Abstract—An attempt is made to explain the single and double inflexions on the daylight *F*-trace in terms of the Appleton–Lyon theory of the "height-lag." Two hypothetical constructs are presented. A procedure for empirically verifying and evaluating either hypothesis is indicated.

Introduction

In a symposium on ionospheric results of solar eclipses under the chairmanship of Ratcliffe (1956), the phenomenon of the $F1\frac{1}{2}$-layer was mentioned and recommended for further study. With this end in view, the related and more regular phenomenon of inflexions on the *F*-trace is undertaken in the present study. It seems that the conditions that normally produce an *F*-trace with a double inflexion on ordinary days are essentially the same as those that produce the $F1\frac{1}{2}$-layer during an eclipse, with the difference that the eclipse intensifies those conditions. At any rate, the subject of *F*-inflexions is sufficiently important by itself to merit a close study.

Appleton and Lyon (1955) have shown that some departures from the Chapman analysis may be accounted for by taking the "height-lag" into account. The present study seeks a connexion between this height-lag and the inflexions on the *F*-trace.

Survey of Data

Selected for this pilot study are 936 ionograms taken at Baguio from 0500 to 1400 hours of each day for 26 successive days of 1956, 26 April to 22 May. This period of the year is selected for convenience and because its solar conditions simulate the summer-equatorial conditions closely enough. The year 1956 is selected because it is about half-way up on the rising slope of the sunspot cycle.

Attention is focused on *first* appearances of single and of double inflexions on the *F*-trace for each day, as well as the general location of the inflexions (low, middle, high) when they first appear. The results are graphically tabulated in Fig. 1.

The general trend shows that between 7 and 9 o'clock, there is only one inflexion, and this is found in the lower part of the *F*-trace; between 9 and 11 o'clock, a second inflexion appears in the upper part. Rare cases of triple inflexions are ignored in the analysis. Sometimes, a large single inflexion in the middle of the trace may also be interpreted as a merging of the upper and lower inflexions.

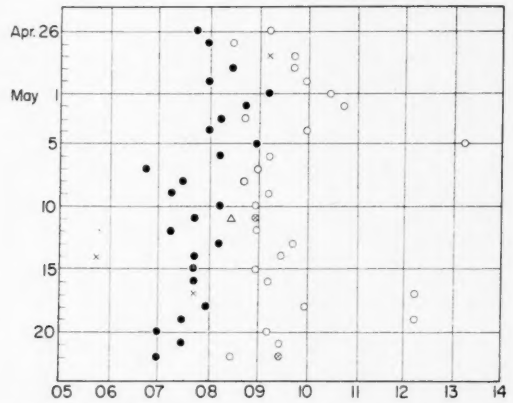


Fig. 1. Time and location of first appearances of inflexions on the *F*-trace: ● single: low; × single: middle; ○ double: high, low; △ double: both low; ⊗ triple: high, middle, low.

VOL. 16 1959

Analysis

Let the coefficient of apparent electron-loss in the *F*-region vary exponentially with height according to the following expression:

$$\alpha = \alpha_0 \exp\left(-\frac{z}{a}\right) \tag{1}$$

where z = true height measured from the bottom of the *F*-region;

$$a = \frac{N_e}{(dN/dz)_e}$$

N_e = electron density at $z = z_e$, with the loss term excluded;

$(dN/dz)_e$ = height-gradient of N at $z = z_e$, with the loss term excluded;

z_e = height of point(s) of inflexion on the *F*-trace.

The assumption and limitations of equation (1) are discussed by Marasigan (1958).

The problem now is to find a function $N(z, \sigma)$, where σ is an arbitrary parameter, such that $a(z_e, \sigma)$ is a curve that yields, for one value of a, either: a single value of z_e for one value of σ; or two values of z_e for another value of σ. The parameter σ is thus seen to control the number of inflexions on the *F*-trace.

To this problem, two simple types of solutions are studied, the arc-sine function and the arc-cosine function. In choosing these, no claim can be made for exclusiveness. It is also pertinent here to recall the need for caution mentioned by Kelso (1954b) about interpreting

$h'(f)$ curves. This warning arose in connexion with his very interesting and instructive study of another arc-cosine model. The following definitions are given:

$$\text{arc-sine function:}\quad N = N_m\left(\frac{2}{\pi}\sin^{-1} z\right)\sigma \tag{2}$$

$$\text{arc-cosine function:}\quad N = N_m\left(1 - \frac{2}{\pi}\cos^{-1} z\right)\sigma \tag{3}$$

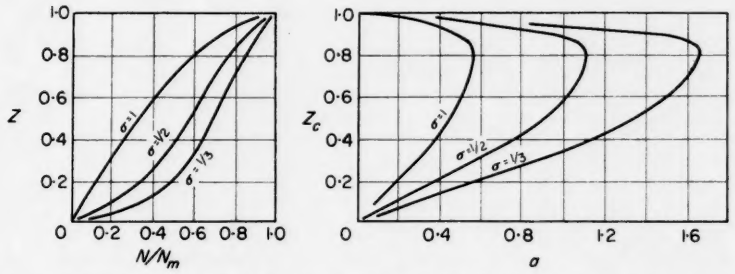


Fig. 2. The curves on the left side are the "lossless distribution" curves for various values of the height-lag γ if $\sigma = \sigma(\gamma)$. The curves on the right side are the corresponding relations between the inflexion-heights Z_c and the constant a of equation (1).

Correspondingly, the a curves are:

$$\text{arc-sine model:}\quad a = \frac{1}{\sigma}\sqrt{(1 - z_c^2)}\,\sin^{-1} z_c \tag{4}$$

$$\text{arc-cosine model:}\quad a = \frac{\pi}{2\sigma}\sqrt{(1 - z_c^2)}\left(1 - \frac{2}{\pi}\cos^{-1} z_c\right) \tag{5}$$

In each case, let $N_m = 1$, $z = 0$ at the bottom of the F-region, and $z = 1$ at the $F2$ peak level. The plots of equations (2) and (4) are given in Fig. 2. The plots of equations (3) and (5) are not given since they resemble those in Fig. 2 except for slight differences in magnitude.

Discussion

Appleton and Lyon (1955) have shown that the ion-density peak lags behind the production peak, say, by a quantity γ measured positively downwards, decreasing with time during the daylight hours. Now if, in equations (2)–(5), $\sigma = \sigma(\gamma)$, it becomes possible to explain the presence of one, and then two, inflexions on the F-trace by such a diurnal variation of the height-lag γ.

A study of the arc-sine function suggests that the larger the height-lag, the smaller the value of σ; for a fixed value of a, this condition results in a tendency of the upper inflexion z_c to rise so close to the $F2$ penetration frequency that it is hardly visible on the F-trace, and only a single inflexion in the lower part of the F-trace is seen. Conversely, the smaller the height-lag (as noon approaches), the larger the value of σ, and the upper inflexion appears, thus accounting for an F-trace with two inflexions. The arc-cosine model behaves similarly except for slight differences in magnitude.

During a solar eclipse at noon, the loss in intensity of solar radiation elevates the production peak so that γ is small, zero or even negative. Thus σ becomes large and calls for the appearance of an $F1\frac{1}{2}$-layer. This is verified at several stations (Hennessey and Torres, 1956; Ratcliffe, 1956; Minnis, 1959).

The problem of a varying scale-height is also related to this discussion. From NICOLET (1954) it can be shown that

$$a = \frac{1}{1 + dH/dg}$$

Thus, with the aid of tables similar to those on the right side of Fig. 2, the height(s) of inflexion(s) on the F-trace can also be used as a rough approximation or empirical check of temperature–gradients in the F-region. But for such a purpose, it will be necessary to include the effects of the transport term. According to MARTYN (1959), both vertical and horizontal transport play a large part in the dynamics of the F-region at locations near the magnetic equator. (When the transport term is too large to be ignored, equation (1) has to be modified accordingly. The procedure for deriving this modification is beyond the scope of this present study.)

SUGGESTIONS FOR FURTHER STUDY

As a result of fruitful studies of BUDDEN (1955), KELSO (1954a, 1954b), SCHMERLING (1958), WRIGHT and NORTON (1959), and others, it is now possible to derive $N(z)$ curves from $h'(f)$ traces on a large scale without too much difficulty. Thus there will soon be sufficient data on true-heights of inflexions with which to evaluate the present theory, provided that the reduction method is shown to be sufficiently accurate, at inflexion frequencies. In any case, it is expected that this inflexion–height hypothesis, in its present form, will undergo modifications—to an extent that still remains to be seen.

Acknowledgements—This study was made possible through the generous assistance of the Reverend J. J. HENNESSEY, S.J., Director of the Manila Observatory. The writer was given free access to all the data and all the literature personally gathered by Father HENNESSEY for the past seven years of ionospheric research.

V. MARASIGAN, S.J.

Manila Observatory
Baguio, Philippines

VOL.
16
1959

REFERENCES

APPLETON E. V. and LYON A. J.	1955	*Physics of the Ionosphere* p. 20. Physical Society, London.
BUDDEN K. G.	1955	*Physics of the Ionosphere* p. 332. Physical Society, London.
HENNESSEY J. J. and TORRES J.	1956	*Solar Eclipses and the Ionosphere* (Edited by BEYNON W. J. G. and BROWN G. J.) p. 65. Pergamon Press, London.
KELSO J. M.	1954a	*J. Atmosph. Terr. Phys.* **5**, 11.
KELSO J. M.	1954b	*J. Atmosph. Terr. Phys.* **5**, 117.
MARASIGAN V.	1958	*J. Atmosph. Terr. Phys.* **13**, 107.
MARTYN D. F.	1959	*Proc. Inst. Radio Engrs., N.Y.* **47**, 147.
MINNIS C. M.	1959	*J. Atmosph. Terr. Phys.* **13**, 346.
NICOLET M.	1954	*The Earth as a Planet* (Edited by KUIPER G. P.) p. 644. Chicago University Press.
RATCLIFFE J. A.	1956	*Solar Eclipses and the Ionosphere* (Edited by BEYNON W. J. G. and BROWN G. J.) p. 147. Pergamon Press, London.
SCHMERLING E. R.	1958	*J. Atmosph. Terr. Phys.* **12**, 8.
WRIGHT J. W. and NORTON R. B.	1959	Private communication.

The gyro-frequency in the *E*-layer above Slough, England

(*Received* 24 *June* 1959)

Abstract—Measurements of the gyro-frequency, f_H, in the *E*-layer from the separation of f_0E and f_xE at Slough give $f_H = 1{\cdot}236 \pm 0{\cdot}015$ Mc/s. The calculated value is $f_H = 1{\cdot}27$ Mc/s. The difference is consistent with that found by Scott using $f_0E - f_xE$ measured at high latitudes.

THE day-time ionograms obtained at Slough during June and July 1954 have been analysed to give accurate values of the separation of the ordinary, f_0, and extra-ordinary, f_x, critical frequencies of the *E*- and E_s-layers. This analysis formed part of an investigation of the mode of reflection in the *E*-layer reported elsewhere (PIGGOTT and BHATTACHARYYA, 1959). The value of the gyro-frequency, f_H, was deduced using the equation:

$$f_0^2 = f_x^2 - f_x f_H \tag{1}$$

The observations were divided into three groups:

(a) those taken between 2330 and 0730 GMT,
(b) those taken between 0730 and 1630 GMT,
(c) those taken between 1630 and 2330 GMT.

No systematic differences in the average values of f_H for the three groups could be detected though the standard deviations were slightly less for group (b) than for the other two groups. In sixty-eight ionograms, all in group (b), both f_xE_s and f_0E_s could be read accurately, and the separations found were indistinguishable from that deduced from f_0E and f_xE for the same period. In almost all these cases $f_0E_s - f_0E$ was near 0·6 Mc/s.

The mean and median values of f_H in the *E*-region above Slough found from all 351 pairs of measurements of ordinary and extraordinary wave critical frequencies are $1{\cdot}23_7$ Mc/s and $1{\cdot}23_5$ Mc/s, respectively. It is unlikely that the probable error, after allowing for systematic experimental error and layer tilts, can exceed $\pm 0{\cdot}015$ Mc/s.

This value may be compared with the ground level value, 1·33 Mc/s (adopting $F = 0{\cdot}475$) and the value 1·27 Mc/s computed from it for a height of 110 km by assuming a mean radius of the earth of 6366 km. A similar slight difference between computed and observed values of f_H was found by SCOTT (1951) who ascribed it to the presence of an appreciable ratio of ions to electrons. This explanation involves very serious theoretical difficulties (MASSEY and BATES, 1946) which have not been resolved. The only alternative appears to be that the earth's magnetic field decreases upwards more rapidly than would be expected from the inverse cube law. Rocket evidence (SINGER *et al.*, 1951; CAHILL and VAN ALLEN, 1958), however, shows that the inverse cube law is correct up to about 95 km and that the decrease associated with penetrating the level of the diurnal current system is only of the order 100 γ. Our discrepancy is of the order 1000 γ.

It is easy to show that, in both sets of data, the apparent value of the magnetic field is smaller than expected by approximately the same amount.

* Official communication.

If we denote the gyro-frequency deduced from equation (1) by f_H and that calculated from the dipole field as $f_H{}^1$, the number of ions per electron needed to cause the discrepancy is essentially proportional to $f_H - f_H{}^1$. If there are λ ions per electron with mean mass μ times the electron mass, SCOTT's (1951) equation for λ/μ can be arranged in the form:

$$\lambda = \mu(f_H{}^1 - f_H) \cdot f_x/(f_x - f_H{}^1) \tag{2}$$

Adopting, for comparison purposes, SCOTT's value $\mu = 4{\cdot}33 \times 10^4$ corresponding to ions of mass 24, we get:

$$\lambda = (2{\cdot}2 \pm 0{\cdot}8) \times 10^3$$

The weighted mean of SCOTT's most accurate determinations of λ, using the difference $f_0E - f_zE$, is $\lambda = 1{\cdot}4 \times 10^3$. The agreement is surprisingly good, indicating that $(f_H{}^1 - f_H)$ is similar in both cases.

Apart from the great theoretical difficulties in accepting values of λ of this order, it is improbable that such high ion to electron ratios could be missed in the rocket experiments (MASSEY, 1959). We are, therefore, forced to suspect that there may be an approximation, of the order of 3 per cent, in the derivation of equation (1).

Further determination of $f_H{}^1 - f_H$ using the intensive measurements made during the IGY would appear to be desirable to clear up this outstanding difficulty.

Acknowledgement—The work described in the paper was carried out as part of the programme of the Radio Research Board and is published by permission of the Director of Radio Research of the Department of Scientific Industrial Research.

W. R. PIGGOTT

DSIR Radio Research Station, Slough

VOL. 16 1959

REFERENCES

BATES D. R. and MASSEY H. S. W.	1946	*Proc. Roy. Soc.* A **187,** 261.
CAHILL, L. J. and VAN ALLEN J. A.	1958	*J. Geophys. Res.* **63,** 270.
MASSEY H. S. W.	1959	Private communication.
PIGGOTT W. R. and BHATTACHARYYA J.	1959	*J. Atmosph. Terr. Phys.* In press
SCOTT C. W.	1951	*J. Geophys. Res.* **56,** 1.
SINGER S. F., MAPLE E. and BOWEN W. A.	1951	*J. Geophys. Res.* **56,** 265.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 199 to 204. Pergamon Press Ltd. Printed in Northern Ireland

Indices of geomagnetic activity

of the Observatories at HARTLAND (*Ha*), ESKDALEMUIR (*Es*), LERWICK (*Le*) December 1958–June 1959

The figures given on pages 199 to 204 represent the K-indices for three-hour intervals, beginning with 00–03 hr for the first and ending with 21–24 hr for the eighth figure.

December 1958

Day	*Ha* Range for $K = 9 : 500\gamma$			*Es* Range for $K = 9 : 750\gamma$			*Le* Range for $K = 9 : 1000\gamma$			*Day*
	K-Indices		*Sum*	*K-Indices*		*Sum*	*K-Indices*		*Sum*	
1	0001	1110	4	0001	1211	6	0001	1111	5	1
2	1012	3355	20	1113	4455	24	2112	3555	24	2
3	3211	0000	7	3312	1110	12	3211	1111	11	3
4	3234	4565	32	3234	4566	33	2123	4667	31	4
5	6422	1112	19	6432	2213	23	8432	2213	25	5
6	1013	3313	15	1112	3323	16	1112	3323	16	6
7	1122	1101	9	2221	1201	11	1221	1101	9	7
8	0002	2243	13	0012	2343	15	0001	2343	13	8
9	3232	2101	14	3232	2211	16	3332	1111	15	9
10	0121	1000	5	1221	1100	8	1111	0001	5	10
11	1112	2114	13	1212	2114	14	1112	2114	13	11
12	2101	3113	12	2201	2222	13	2201	1113	11	12
13	5532	5565	36	4423	4565	33	5421	4775	35	13
14	3212	4344	23	3212	3344	22	4121	3345	23	14
15	1102	1333	14	1113	2232	15	2112	1322	14	15
16	3323	3220	18	3323	3221	19	2312	2211	14	16
17	1002	2564	20	1112	2466	23	1111	1467	22	17
18	6522	3123	24	6422	3213	23	8722	3223	29	18
19	4323	3332	23	4323	3332	23	5533	3332	27	19
20	2333	3333	23	2332	2333	21	3332	2233	21	20
21	2122	2131	14	2222	1122	14	2121	1221	12	21
22	3011	2133	14	3111	2232	15	3111	2233	16	22
23	0012	3342	15	1122	3332	17	1122	2432	17	23
24	2112	1012	10	2122	2112	13	2222	1112	13	24
25	0001	1113	7	0001	1123	8	0011	2122	9	25
26	2111	3443	19	2111	3343	18	1112	2333	16	26
27	2222	3301	15	2222	3311	16	2222	2211	14	27
28	3223	3332	21	3232	3332	21	3222	2232	18	28
29	2112	3122	14	2211	2133	15	3111	2133	15	29
30	2102	3332	16	2121	3322	16	2112	2233	16	30
31	2102	2120	10	2111	2121	11	2111	2121	11	31

January 1959

Day	*Ha* *Range for* $K = 9 : 500\gamma$			*Es* *Range for* $K = 9 : 750\gamma$			*Le* *Range for* $K = 9 : 1000\gamma$			*Day*
	K-Indices		*Sum*	*K-Indices*		*Sum*	*K-Indices*		*Sum*	
1	0001	1000	2	0010	1000	2	0011	1111	6	1
2	0002	2110	6	0002	1111	6	1011	1111	7	2
3	2123	3110	12	2222	3121	15	2222	2121	14	3
4	1012	2332	14	1012	2331	13	1112	2332	15	4
5	2103	4455	24	2103	3446	23	1112	2367	25	5
6	4222	3444	25	4232	3343	24	4322	2344	24	6
7	3332	2244	23	4332	2244	24	4332	1255	25	7
8	4222	2345	24	4222	2345	24	4322	2345	25	8
9	4113	3545	26	4222	3746	30	5212	3957	34	9
10	5333	3354	29	4333	4354	29	6334	4575	37	10
11	4113	2133	18	4212	2133	18	6211	2224	20	11
12	3213	2333	20	4212	2233	19	4212	2232	18	12
13	1112	2232	14	1121	1231	12	2122	1221	13	13
14	0022	2222	12	0022	3223	14	1121	2113	12	14
15	2321	1000	9	2321	1121	13	2332	2111	15	15
16	1012	3344	18	1003	3433	17	1012	3334	17	16
17	2322	3332	20	2332	3332	21	3322	3331	20	17
18	3322	2434	23	3322	2434	23	3322	2433	22	18
19	2222	2142	17	2232	2131	16	1223	2132	16	19
20	0002	1123	9	1011	1122	9	1013	1123	12	20
21	0011	1121	7	1011	2120	8	1011	1111	7	21
22	1121	1122	11	1122	1222	13	1121	1122	11	22
23	3102	3021	12	3112	3221	15	3212	2121	14	23
24	1111	0100	5	2221	1100	9	1111	1111	8	24
25	1103	4453	21	2103	4453	22	1013	3462	20	25
26	1313	4310	16	1313	4421	19	1213	4321	17	26
27	0112	2222	12	0123	2223	15	1112	2223	14	27
28	3110	2222	13	3223	3222	19	3212	2232	17	28
29	1113	3344	20	2112	3354	21	1112	2444	19	29
30	3222	3220	16	3322	2221	17	3212	2230	15	30
31	3332	2122	18	3333	2222	20	3332	1222	18	31

February 1959

1	2312	3222	17	2222	3232	18	2221	2232	16	1
2	1333	3334	23	2223	3334	22	1333	3334	23	2
3	4213	3335	24	3223	3334	23	3223	3335	24	3
4	4332	4354	28	4333	4445	30	4323	3457	31	4
5	3322	2344	23	3333	3333	24	3333	3444	27	5
6	4212	3344	23	4213	3343	23	3323	2344	24	6
7	3212	3232	18	2222	3231	17	3212	2232	17	7
8	0013	3234	16	2122	3334	20	1122	3235	18	8
9	4442	3112	21	4432	3223	23	4433	2222	22	9
10	2001	1112	8	2102	1223	13	2101	1112	9	10
11	1335	5424	27	1335	5434	28	1324	4426	26	11

VOL. 16 1959

February 1959 (*Continued*)

Day	*Ha* Range for $K = 9 : 500\gamma$			*Es* Range for $K = 9 : 750\gamma$			*Le* Range for $K = 9 : 1000\gamma$			*Day*
	K-Indices		*Sum*	*K-Indices*		*Sum*	*K-Indices*		*Sum*	
12	4423	3334	26	3334	3333	25	4433	2235	26	12
13	3113	3344	22	3113	3444	23	4112	3444	23	13
14	3324	5433	27	3233	4533	26	3333	4553	29	14
15	3443	3441	26	3443	3441	26	3553	3441	28	15
16	5542	4565	36	4543	4576	38	5542	4677	40	16
17	5432	1020	17	5432	2221	21	6542	2222	25	17
18	1001	1211	7	2111	1211	10	2111	1222	12	18
19	2332	2120	15	3332	2221	18	2332	2220	16	19
20	0012	1100	5	0112	2111	9	0012	1111	7	20
21	0022	2122	11	1022	2122	12	0012	2112	9	21
22	2423	2313	20	2333	2323	21	2333	2213	19	22
23	4333	2000	15	4232	1100	13	4333	2110	17	23
24	0012	2112	9	0012	2212	10	0012	2111	8	24
25	3454	4754	36	3444	5655	36	2344	6766	38	25
26	5453	4444	33	4354	4454	33	5453	4454	34	26
27	3322	4653	28	3222	3543	24	3322	3745	29	27
28	4423	4564	32	4433	4464	32	4423	4476	34	28

March 1959

1	4443	3545	32	4443	4544	32	5543	5555	37	1
2	3343	3454	29	3343	4454	30	4342	4464	31	2
3	3323	2344	24	3323	3444	26	3333	3345	27	3
4	3222	2242	19	3233	3342	23	4222	2232	19	4
5	3222	1243	19	3222	2243	20	3221	2343	20	5
6	3310	0020	9	3211	2111	12	3211	2120	12	6
7	0001	1233	10	1002	1233	12	1112	2434	18	7
8	4311	2111	14	4312	1212	16	4321	1212	16	8
9	1111	1100	6	2111	1111	9	2111	1110	8	9
10	0021	0000	3	0022	2101	8	0022	1100	6	10
11	0010	2021	6	1011	2121	9	1011	1111	7	11
12	3333	3330	21	3333	4332	24	3223	3331	20	12
13	1222	3114	16	2223	3324	21	1222	3214	17	13
14	0123	2112	12	1133	2212	15	1222	2112	13	14
15	3221	2002	12	2212	2212	14	2212	1102	11	15
16	3010	1111	8	2112	2121	12	2012	1111	9	16
17	1220	1112	10	2222	2222	16	2122	1212	13	17
18	0111	1204	10	0102	2214	12	0211	2213	12	18
19	0012	2220	9	1122	3221	14	1112	2221	12	19
20	1211	2001	8	1222	2211	13	1221	2211	12	20
21	1012	2211	10	1112	2311	12	1112	2221	12	21
22	0001	2112	7	1112	3212	13	1122	2212	13	22

March 1959 (*Continued*)

Day	*Ha* Range for $K = 9 : 500\gamma$		Sum	*Es* Range for $K = 9 : 750\gamma$		Sum	*Le* Range for $K = 9 : 1000\gamma$		Sum	Day
	K-Indices			*K-Indices*			*K-Indices*			
23	2321	1232	16	2321	2332	18	2321	1222	15	23
24	2123	2221	15	2113	3322	17	2123	3321	17	24
25	2522	4543	27	2422	5543	27	1422	4462	25	25
26	2035	6556	32	2135	7556	34	1134	7468	34	26
27	6574	6875	48	6565	7775	48	7785	7798	58	27
28	2434	4566	34	3434	5678	40	5433	6797	44	28
29	5554	4652	36	6555	5642	38	7654	6642	40	29
30	2334	2222	20	2334	3223	22	4434	2233	25	30
31	2222	3244	21	2222	3345	23	3322	3345	25	31

April 1959

1	3330	1110	12	3331	1110	13	4341	1110	15	1
2	0012	1122	9	0122	2223	14	1022	2232	14	2
3	3322	2203	17	3323	2313	20	3322	2213	18	3
4	3221	0211	12	3221	1211	13	2222	2211	14	4
5	2111	1210	9	2122	2221	14	2022	2211	12	5
6	2011	1222	11	2021	2232	14	2122	2232	16	6
7	1112	2243	16	2123	2243	19	1112	2243	16	7
8	2332	3324	22	3333	3324	24	3322	3446	27	8
9	5543	2254	30	5533	2354	30	7532	1244	28	9
10	4466	6643	39	4476	6753	42	6566	7763	46	10
11	3223	3335	24	3223	3335	24	3233	2336	25	11
12	3323	2102	16	3323	2212	18	3222	2312	17	12
13	3122	3232	18	3222	3331	19	2121	2222	14	13
14	3221	2313	17	3222	2323	19	3322	2214	19	14
15	3121	2221	14	3222	1232	17	5221	1231	17	15
16	2113	1121	12	2213	2221	15	2112	2221	13	16
17	0022	2122	11	1123	3222	16	1122	3221	14	17
18	3001	2220	10	3012	2220	12	2111	1220	10	18
19	0001	2312	9	0112	2212	11	1011	2312	11	19
20	0112	2211	10	0122	3211	12	1202	2211	11	20
21	2222	3310	15	3222	2311	16	2222	2320	15	21
22	0001	1100	3	0102	2210	8	1112	2101	9	22
23	0003	4465	22	0004	5577	28	0113	4666	27	23
24	3343	3414	25	3354	4424	29	4343	4434	29	24
25	3232	3433	23	3332	3433	24	3332	3544	27	25
26	4423	2334	25	3323	2334	23	4422	2444	26	26
27	2231	2324	19	2232	3323	20	4232	3334	24	27
28	1122	2334	18	1122	3334	19	2222	2336	22	28
29	3322	4543	26	3322	4542	25	4322	4653	29	29
30	1122	4552	22	2223	4552	25	2323	5543	27	30

VOL. 16 1959

May 1959

Day	*Ha* Range for $K = 9 : 500\gamma$		*Es* Range for $K = 9 : 750\gamma$		*Le* Range for $K = 9 : 1000\gamma$		*Day*
	K-Indices	*Sum*	*K-Indices*	*Sum*	*K-Indices*	*Sum*	
1	1302 2232	15	1312 3222	16	2322 2222	17	1
2	2221 1122	13	2233 2122	17	3222 1122	15	2
3	1011 2224	13	1101 2224	13	1011 1214	11	3
4	3122 2234	19	3122 3334	21	3122 2225	19	4
5	5423 3210	20	6434 4321	27	6623 3221	25	5
6	0011 1200	5	0102 2311	10	0011 1200	5	6
7	1021 1113	10	1112 2223	14	1012 2113	11	7
8	3543 4423	28	3443 4433	28	3544 3322	26	8
9	3223 4423	23	3223 4423	23	3323 2442	23	9
10	3233 3432	23	3233 3442	24	3233 4423	24	10
11	3233 3435	26	3233 4435	27	2233 3434	24	11
12	6555 6844	43	6554 6855	44	5754 5866	46	12
13	1231 3333	19	2232 3434	23	1232 3333	20	13
14	1112 2100	8	2212 3221	15	1211 3211	12	14
15	2124 5654	29	2224 5654	30	2224 5656	32	15
16	5543 2314	27	5553 2324	29	7653 2224	31	16
17	1122 3422	17	2222 3433	21	2123 3432	20	17
18	3331 3345	25	3332 3444	26	4432 2445	28	18
19	3222 2332	19	3323 3333	23	3323 3333	23	19
20	3122 2343	20	3223 2332	20	3212 2332	18	20
21	2322 2333	20	2323 2343	22	3322 2433	22	21
22	3132 2332	19	3233 2433	23	4322 3333	23	22
23	3322 2222	18	3323 2332	21	3323 2332	21	23
24	3354 4446	33	3344 4446	32	4344 4448	35	24
25	5333 3422	25	5333 3433	27	8333 3433	30	25
26	3102 2231	14	3112 2332	17	4112 2332	18	26
27	2111 1021	9	2212 2111	12	3212 1111	12	27
28	0011 2111	7	1112 2221	12	1111 1011	7	28
29	0000 0111	3	1011 1122	9	1112 0111	8	29
30	1002 4320	12	1002 4322	14	2112 3221	14	30
31	2322 4344	24	2322 5344	25	1322 4344	23	31

June 1959

1	2220 2222	14	3222 2232	18	3221 1223	16	1
2	2223 4343	23	2233 4343	24	2223 4345	25	2
3	3323 4322	22	3333 4422	24	4423 4331	24	3
4	3233 3334	24	3343 3334	26	4333 3335	27	4
5	4222 3332	21	4323 4332	24	5333 3322	24	5
6	1223 3332	19	2223 3433	22	1123 3433	20	6
7	2321 2331	17	3322 3332	21	3321 2321	17	7
8	2222 3433	21	1122 3433	19	1122 3322	16	8
9	3333 3332	23	3333 3332	23	2333 3332	22	9
10	3422 2212	18	3322 3222	19	4332 2211	18	10
11	1124 5431	21	1124 6422	22	1114 5431	20	11

June 1959 (*Continued*)

Day	*Ha* Range for $K = 9 : 500\gamma$			*Es* Range for $K = 9 : 750\gamma$			*Le* Range for $K = 9 : 1000\gamma$			*Day*
	K-Indices		*Sum*	*K-Indices*		*Sum*	*K-Indices*		*Sum*	
12	1011	2331	12	1121	2432	16	1121	2331	14	12
13	1101	2111	8	1112	2221	12	1121	1211	10	13
14	0222	3313	16	1213	4323	19	1213	3322	17	14
15	1332	2111	14	2233	3121	17	2332	2211	16	15
16	1011	2122	10	2122	3232	17	2222	3222	17	16
17	0000	2233	10	0011	2233	12	1102	2223	13	17
18	2222	3322	18	2221	4321	17	2212	3322	17	18
19	1222	3121	14	1322	3232	18	1222	3221	15	19
20	3112	3132	16	2112	3233	17	3222	3222	18	20
21	2021	2332	15	2121	3332	17	2111	2322	14	21
22	1223	3333	20	1123	3333	19	1123	3322	17	22
23	2221	1443	19	3122	2444	22	2222	1435	21	23
24	4410	0232	16	4413	1342	22	6632	1343	28	24
25	2112	1213	13	2122	2223	16	2112	2213	14	25
26	2332	2320	17	3342	3331	22	3332	3331	21	26
27	0022	5634	22	1022	6635	25	1112	6555	26	27
28	4433	4333	27	4432	5432	27	5543	5544	35	28
29	3355	4325	30	3355	4445	33	4444	4566	37	29
30	4234	5432	27	4234	5433	28	6434	6543	35	30

VOL. 16 1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 205 to 219. Pergamon Press Ltd. Printed in Northern Ireland

Spectroscopic observations of the great aurora of 10 February 1958—I

Abnormal vibration of N_2^+

K. C. CLARK* and A. E. BELON
Geophysical Institute, University of Alaska, College, Alaska

(*Received* 16 *December* 1958)

Abstract—Auroral spectrograms taken with the IGY patrol spectrograph at College, Alaska, during the great aurora of 10–11 February 1958, show high vibrational excitation in the $B(^2\Sigma_u^+)$ state of N_2^+, as indicated by the enhanced intensities of the 4652 Å (1,3) and 4600 Å (2,4) bands relative to the 4709 Å (0,2) band of the first negative system. The grating spectrograms presented consist of 18 successive exposures throughout the night from 0308 to 1529 hours UT 11 February and resolve the magnetic meridian from North to South along the length of the slit image. Intensity ratios are compared with those of synthetic spectra at the same 8 Å resolution calculated assuming Boltzmann distributions at various excitation temperatures of vibration and of rotation. The excitation can in this manner be roughly characterized by $T(\text{vib}) = 4000°\text{K}$ and $T(\text{rot}) = 2500°\text{K}$. The enhancement is frequently so strong as to reverse the normally decreasing order of intensities of the (0,2) and (1,3) bands. Unusually intense emission from hydrogen is frequently present, and the measured Doppler shifts of H_α indicate velocities up to at least 800 km/sec. These auroral measurements lend observational support to the concept that charge exchange becomes a major process in great aurorae.

INTRODUCTION

AURORAL observations of high vibration in N_2^+ are uncommon. BATES (1949, 1954) discusses the available data. He calculates vibration temperatures of about 2000°K for the high latitude aurorae of VEGARD and KVIFTE (1945), and about 8000°K for the low latitude observations of BARBIER (1947) if electron excitation is assumed. RAYLEIGH's (1922) spectrogram of a low latitude aurora over southern England gives relative intensities of 2, 3, 2 for the (0,2), (1,3), (2,4) sequence of the first negative system of N_2^+. STÖRMER (1939) reports relative intensities of 1·00, 0·97, 0·85 for high sunlit rays. Moderate vibration enhancements are reported by BARBIER and WILLIAMS (1950) in Alaska and in lesser degree by PETRIE and SMALL (1952) in Saskatoon. However, HUNTEN (1955, 1956) recently gives low vibration intensity ratios, indicating excitation by electron impact and leading to uncertain atmospheric temperatures in the range 0–700°K. The normal intensity ratios given by VEGARD (1940) are similarly low.

Laboratory measurements using 6 keV canal rays to excite the negative bands of N_2^+ have been made by SMYTH and ARNOTT (1930), who report large enhancement of vibration and rotation. BRANSCOMB *et al.* (1954) conducted quantitative experiments with protons in nitrogen, and present their resulting measurements of high vibration enhancement in comparison with the known laboratory and auroral results of others. FAN (1956a,b) gives measurements and excellent spectrograms showing the excitation by H^+ and He^+ at 5–450 keV energies in traversing air and other gases. Helium ions and, to a lesser extent, protons produce very strong enhancement of vibration in N_2^+ when their energies are low enough to inhibit direct collision ionization and yet permit charge exchange excitation. CARLETON (1957), using 3 keV protons in N_2, failed to detect any enhancement of rotation temperature and presumably of vibration temperature. ROESLER *et al.*

* On leave from University of Washington, Seattle, 1957–1958.

(1958) also failed to observe any enhanced rotation using protons of 10–30 keV energies in N_2.

In the spectra reported in the present paper, relating to the great type-*A* red aurora of 10–11 February 1958, a very strong continuing enhancement of N_2^+ vibration was observed along with other unusual spectroscopic features. Very intense hydrogen lines were present in the early hours. A discussion of enhanced atomic emissions is given in a second paper (Belon and Clark, 1959), while the present report concerns features mainly related to the vibrational enhancement of the $\Delta v = 2$ sequence of the N_2^+ first negative bands. The eighteen spectrograms taken consecutively throughout the night are in effect multiple, since each gives resolution from North to South horizon along the magnetic meridian. General observations associated with this auroral display are also summarized.

General Description of the Aurora

Chronologically, the first associated event observed at College* was the start of prolonged and intense emission of solar radio noise at 65 Mc/s. This noise is known to correlate with solar flare activity (Dodson *et al.*, 1953; Warwick, 1954), and at times it presages the onset of strong auroral activity. After an interval of 55 hr the following terrestrial effects were noted.

At 1530 hours on 10 February (150° WMT)† there were sudden commencements of magnetic disturbances and of the most intense earth current fluctuations in recent years. No commencement of visible aurora was seen at College, for the time of sunset was 1603 hours. However, auroral all-sky cameras operating in the evening at Delaware, Ohio and Hanover, New Hampshire, showed the start of a faint northern arc at 1526 (150° WMT) which became strong and rayed at 1558 and developed into the night-long display. This aurora was reported seen from as far south as Cuba. Vertical ionosphere soundings in the 1–15 Mc/s range showed full absorption before 1630, during 2250–0045, and after 0730; there was spread-*F* reflection from 1800 to 2000. Auroral radar reflections appeared weakly at 1619–1622 and 1630–1631, becoming strong during 1652–1710 and continuing intermittent until a strong development at 0040–0113. Absorption of 27·6 Mc/s cosmic noise shown by the network of IGY riometer stations indicated that the absorption zone was broadened and shifted southward from the usual auroral zone. In only the auroral zone there was extremely strong day-time cosmic noise absorption during 0800–1700, particularly during 1400–1600. This absorption continued high and unusually constant through the night, showing, as with other stations, a strong increase for an hour or two after midnight.

The visual aurora appeared intense red even in twilight, frequently covered most of the sky, and was unusually diffuse and stable until 2100. From comparative local observations it was judged to have considerably surpassed the great aurora of 23 February, 1956, in duration and intensity. Two or more major arcs could be seen at times from College, and extremely long red rays developed occasionally later in the night. Reliable triangulation of features from records of Alaskan automatic all-sky cameras was in general handicapped by the unusual breadth and diffuseness of identifiable forms. The procedure of ignoring all but indisputable features unfortunately eliminated most of the desirable times, but because of their pertinence the remaining approximate heights are given in Table 1. Two geomagnetic N–S baselines were used. Barrow–Kotzebue, 540 km and Ft. Yukon–College, 210 km. Kotzebue lies geomagnetically 710 km west and only 100 km south of College; in view of the broad range in latitude of the arcs it is not unreasonable to assume that heights of features viewed in the Kotzebue meridian are also typical of those seen in the College meridian.

Because of the breadth of the displays a few of the bases whose heights are measured lie

VOL. 16 1959

* The authors are indebted to Dr. L. Owren, Dr. V. Hessler, Dr. C. G. Little and Mr. H. Leinbach of the Geophysical Institute staff for discussion of the following electric and magnetic effects.

† 150° West Meridian Time (Alaska Standard) will be used consistently in this paper. Thus, 1530 10 February (150° WMT) = 0130 11 February (UT).

Table 1. Heights (km) of measurable features in magnetic meridian. North or South elevation angles at Kotzebue or College are given. Diffuse homogeneous arcs, rayed arcs, and rays are indicated

Time	Place	Form	Base elevation	Base height	Top elevation	Top height
1910	College	D.H.A.			S70	380
1915	Kotzebue	D.H.A.	N 88	340	Z90	620
1935	College	D.H.A.			S75	420
1940	Kotzebue	D.H.A.			N70	370
1948	Kotzebue	D.H.A.	S 80	120	S80	490
2014	Kotzebue	D.H.A.	S 30	150	S60	420
2029	College	H.A.			S40	560
2031	College	H.A.			S45	700
2032	College	H.A.	S 7	140 S	S50	360
2033	College	H.A.	S 8	110 S	S55	300
2034	College	H.A.			S57	310
2035	College	H.A.			S64	370
2036	College	H.A.	S 6	80 S	S67	360
2037	College	H.A.	S 6	110 S	S67	460
2038	College	H.A.			S70	460
2039	College	H.A.			S72	300
2040	College	H.A.			S72	340
2041	College	H.A.			S75	370
2042	College	H.A.			S80	330
2043	Kotzebue	H.A.			S67	240
2043	Kotzebue	R.A.	N 55	190	Z90	520
2105	Kotzebue	R.A.			S80	600
2338	College	H.A.			S40	280
0015	Kotzebue	H.A.	S 78	110	S35	400
0030	Kotzebue	R.A.	S 20	120	S57	620
0032	Kotzebue	R.	N 60	180	Z90	370
0107	Kotzebue	R.	N 10	90	N60	770
0113	Kotzebue	R.A.	N 10	90 S	N55	720
0127	Kotzebue	R.	N 13	90	N80	900
0133	Kotzebue	R.	N 15	110	N80	1100
0147	Kotzebue	R.	N 18	120	N80	850
0224	College	R.A.			S60	390
0225	Kotzebue	R.	N 32	160	N81	800
0302	Kotzebue	R.A.	N 22	110	N80	690
0308	Kotzebue	R.	N 10	100 S	N45	530
0340	Kotzebue	R.	N 10	100	N50	500
0345	Kotzebue	R.	N 5	150	N55	800
0424	Kotzebue	R.	N 12	100	N60	660
0440	College	R.			S73	270
0446	College	R.A.	N 27	150		
0516	College	R.			S70	290
0540	Kotzebue	R.	N 25	100	Z90	700

North or South of the field line through the tops measured; these independent base heights are labelled N or S. The Kotzebue measurements almost all show the correct alignment characteristic of single features. It is seen that in many cases the top heights definitely exceed 500 km, and are sometimes 1000 km. Base heights, which are more frequently subject to ambiguity, vary from the normal 100 km region to values frequently in the 200–300 km range. Some of the highest rays have low bases.

Spectrum Analysis

The IGY patrol spectrograph gives a medium resolution of 8 Å and records successive grating spectra on 16 mm Eastman 103a-F film with a semi-solid Schmidt mirror of $f/0{\cdot}625$ aperture. The all-sky field lens places on the entrance slit an image of the sky extending from North to South along the geomagnetic meridian. The geometric shadow line in the magnetic meridian is shown in Fig. 1 as a function of height and elevation angle for evening and morning hours. Fig. 2 presents chronologically the eighteen patrol spectra obtained during this night together with selected typical College all-sky photographs regularly spaced through the duration of spectrum exposure. The solar continuum is present in the first three spectra. The enhanced oxygen red lines are striking, and even in a short exposure of one minute, the intense H_α emission is recorded.

An additional time-resolved spectrum for the magnetic zenith is given by the Huet $f/0{\cdot}7$ prism spectrograph in routine patrol operation. In this instrument the plate moves perpendicular to the dispersion, and the short slit gives an effective time resolution of 1 hr. This spectrum, kindly furnished by Romick (1959), is shown in Fig. 3 with time and wavelength scales, and gives a very sensitive log of zenith auroral activity.

VOL. 16 1959

The eighteen patrol spectrograms were analysed with a recording microdensitometer in four independent 45° segments of elevation as labelled in Fig. 1, giving a total of seventy-two densitometer traces. Emulsion calibration was made using a known incandescent lamp, step exposures and proper correction for background density. For detailed analysis, nineteen typical spectra were converted to an intensity scale. Narrow atomic lines show essentially no change in half width even at high densities. Therefore the wavelength resolution of the spectrum can be confidently set at 8 Å.

Hydrogen emissions are shown intensely in most of the spectrograms until about 0100 hours and remain detectable thereafter. There is indication of hydrogen in great abundance, of characteristic Doppler broadening and of possibly slower maximum speeds than reported in the more detailed measurements of Meinel (1951). The shift of the H_α line has been determined as a function of elevation angle, although the moderate resolution is not suited to high accuracy. Repeated measurements on spectrograms nos. 2, 3 and 7 were made using a travelling stage microscope. Positions of the peak and of the extremes of recognizable line density were recorded, along with locations of other close atomic lines. The shifted contours of H_α are given in Fig. 4 for each of the three spectra. Since these measurements are intended to give the maximum observed velocity, not a half width, the varying position of the extremes may be due to intensity changes as well as to velocity changes. Near the horizon the peak of H_α agrees well with 6563 Å. The maximum displacement due to Doppler shift occurs near the magnetic zenith,

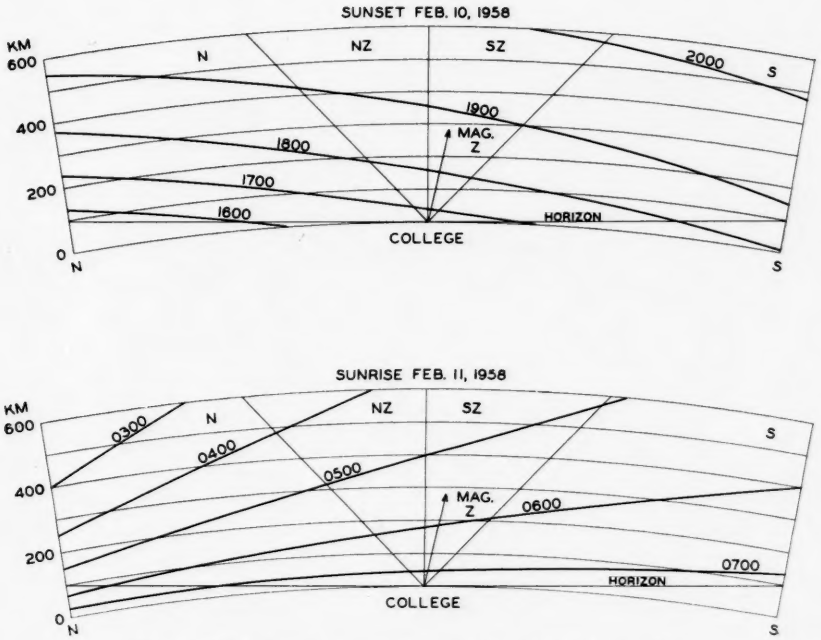
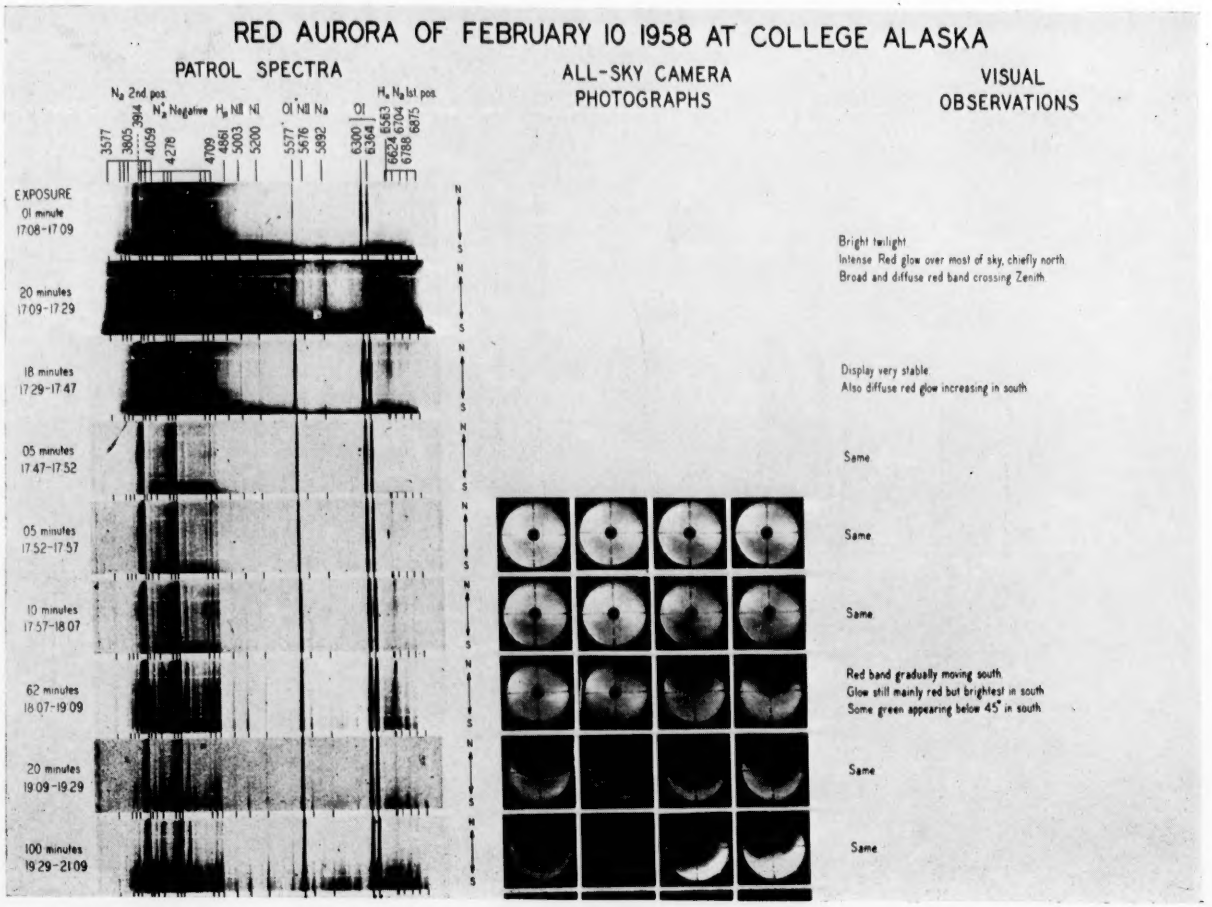


Fig. 1. Shadow line in magnetic meridian at 1 hr intervals.

RED AURORA OF FEBRUARY 10 1958 AT COLLEGE ALASKA



VOL. 16 1959

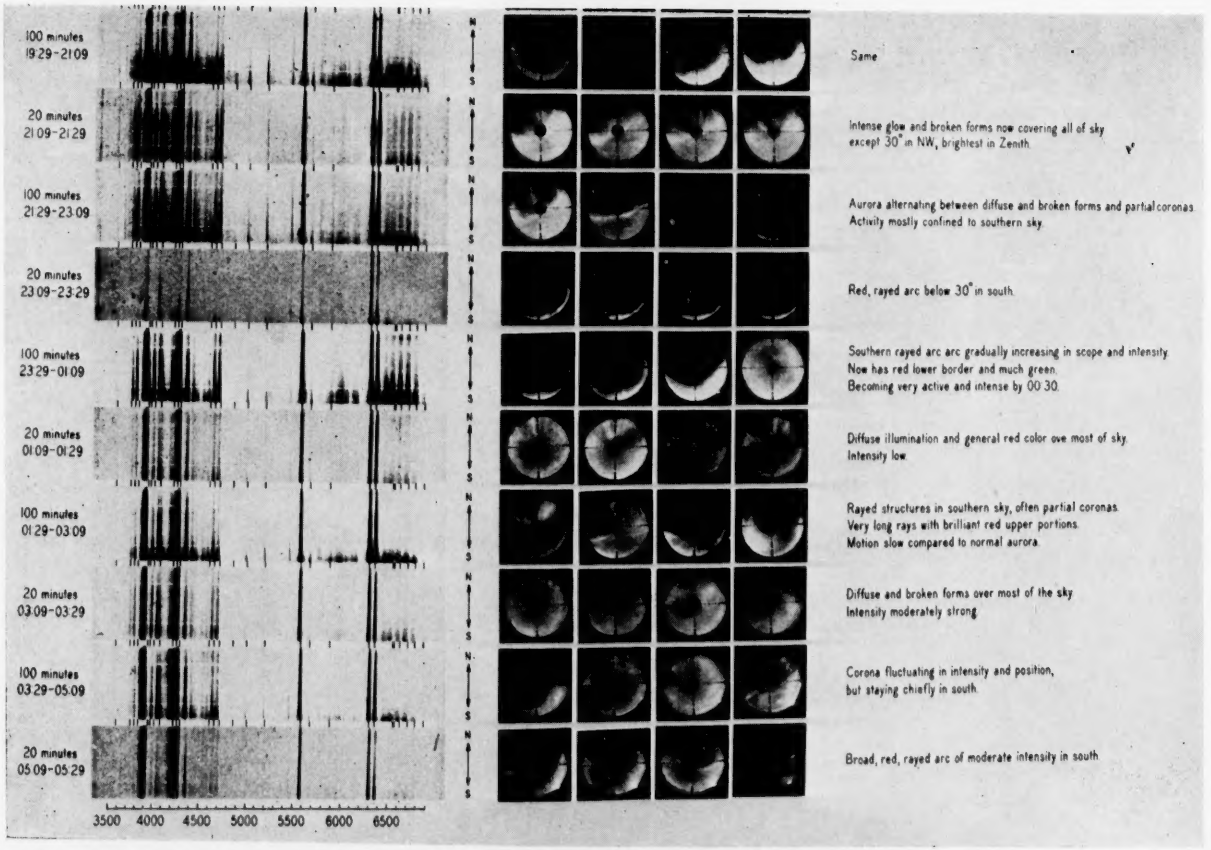

100 minutes 19:29-21:09
20 minutes 21:09-21:29
100 minutes 21:29-23:09
20 minutes 23:09-23:29
100 minutes 23:29-01:09
20 minutes 01:09-01:29
100 minutes 01:29-03:09
20 minutes 03:09-03:29
100 minutes 03:29-05:09
20 minutes 05:09-05:29
3500 4000 4500 5000 5500 6000 6500
Same
Intense glow and broken forms now covering all of sky except 30° in NW, brightest in Zenith.
Aurora alternating between diffuse and broken forms and partial coronas. Activity mostly confined to southern sky.
Red, rayed arc below 30° in south.
Southern rayed arc arc gradually increasing in scope and intensity. Now has red lower border and much green. Becoming very active and intense by 00:30.
Diffuse illumination and general red color ove most of sky. Intensity low.
Rayed structures in southern sky, often partial coronas. Very long rays with brilliant red upper portions. Motion slow compared to normal aurora.
Diffuse and broken forms over most of the sky. Intensity moderately strong.
Corona fluctuating in intensity and position, but staying chiefly in south.
Broad, red, rayed arc of moderate intensity in south.

Fig. 2. Patrol spectrograms, typical all-sky camera photographs and visual observations.

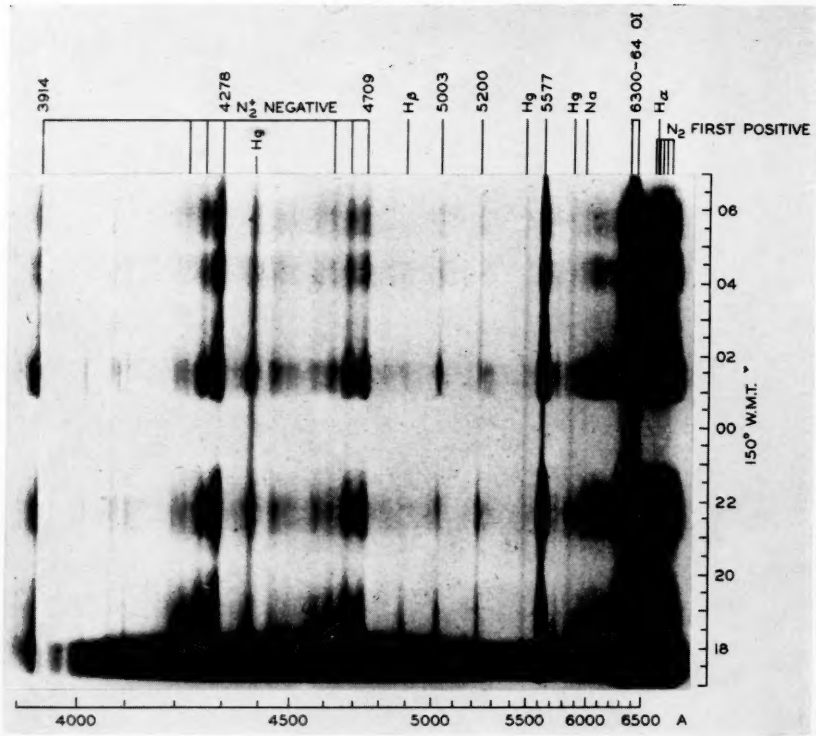


Fig. 3. Time-resolved, zenith spectra taken with moving plate and prism dispersion. Red aurora of February 10, 1958, at College, Alaska.

VOL.
16
1959

which is at S 77° elevation, and amounts to 17 Å, or a velocity of approach of 800 km/sec.

The $\Delta v = 2$ sequence of N_2^+ at 4709 Å (0,2), 4652 Å (1,3), 4600 Å (2,4) was chosen for study because the more intense sequences with heads at 4278 Å ($\Delta v = 1$) and 3914 Å ($\Delta v = 0$) were too heavily exposed for measurement and furthermore are inherently less sensitive to vibration enhancement. The Frank–Condon factors are such that, in the excitation conditions studied, the intensity of

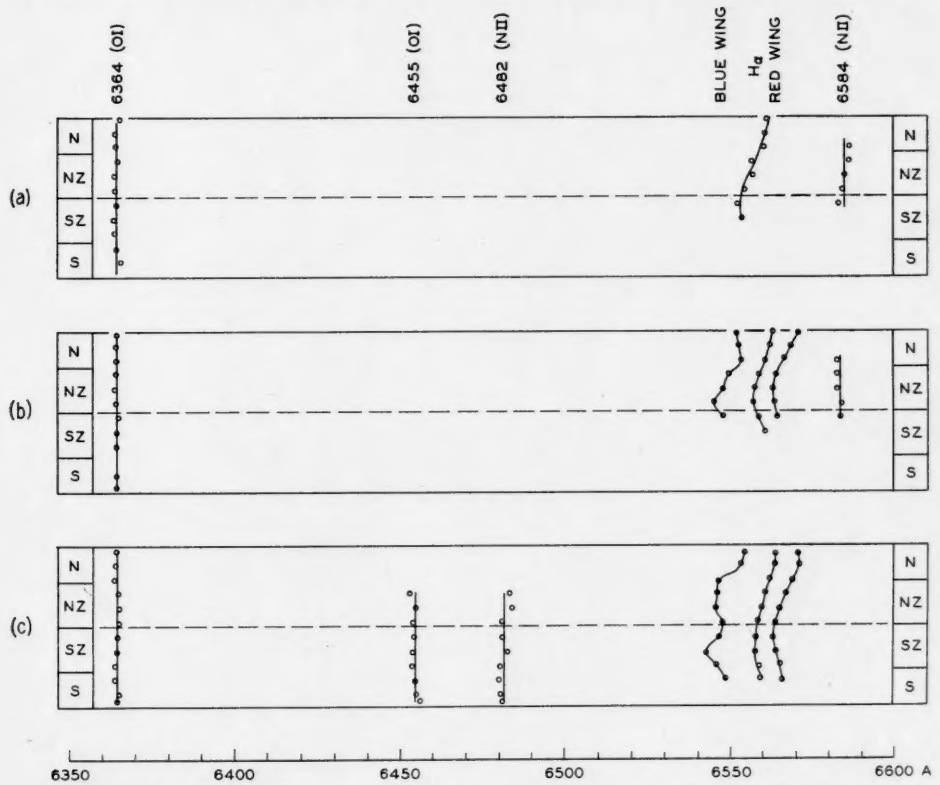


Fig. 4. Observed Doppler shifts of H_α relative to neighbouring wavelengths. Centre line locates peak; side lines show extent of definitely detectable emission. Blue shift of 17 Å indicates velocity of 800 km/sec. (a) Spectrum 2: 1709–1729 hours, (b) Spectrum 3: 1729–1747 hours, (c) Spectrum 7: 1807–1909 hours.

(1,3) not only can show a smaller decrease than usual relative to the (0,2) intensity but can even exceed it; this inversion is visible in Figs. 2 and 3 qualitatively without recourse to emulsion calibration. It persists through much of the night and is not found in any other of the routine College spectrograph exposures taken throughout 1957–1958. Fig. 5 shows comparative densitometer traces of a typical enhancement observed in this red aurora and of the ordinary distribution in a normal intense aurora (100 min exposure, early evening, NZ, 17 February 1958).

It appears certain that other known lines or bands have not made appreciable contributions to the intensities at 4652 Å (1,3) and 4600 Å (2,4). The observed

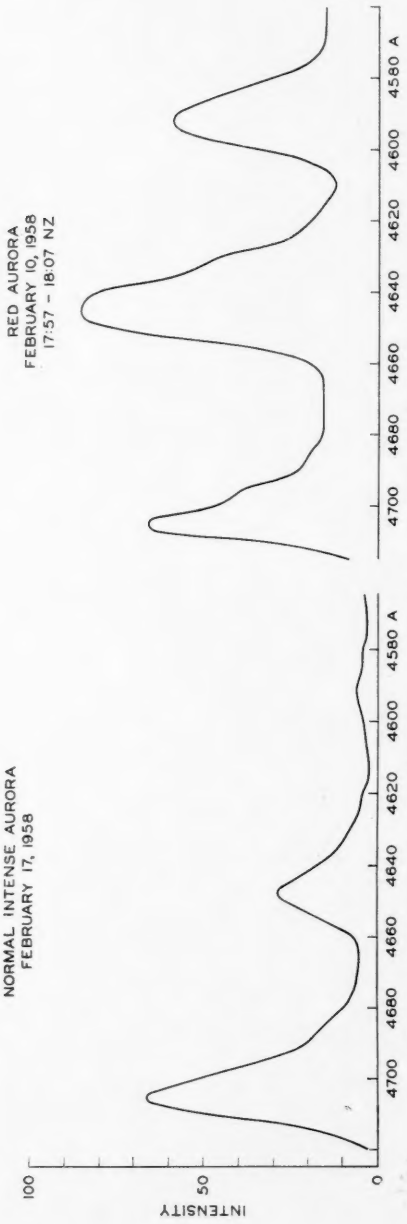


Fig. 5. Enhancement of high vibration in $\Delta v = 2$ sequence at 8 Å resolution. Ordinary distribution of (0,2), (1,3), (2,4) peaks is illustrated by 17 February spectrum.

VOL. 16 1959

half widths of these bands are several times greater than those of known atomic lines of higher densities on the same spectrograms. Some of the total intensity at the (1,3) band observed by VEGARD (1958) at 8 Å resolution and by HUNTEN (1955) at 10–20 Å resolution has been attributed by them to the 4651 Å and the 4639 Å, 4642 Å members, respectively, of the multiplet no. 1 of *OII* ($3s^4P$–$3p^4D^0$). Members of this multiplet can indeed be observed in the present spectra, as will be discussed in Part II, and particularly the 4676 Å member provides a weak wavelength reference; but they contribute negligibly to the band intensity. Contamination by a 4650 Å (4,16) band of the Vegard–Kaplan system, as reported by VEGARD (1958), does not seem likely, inasmuch as this is presumably a weaker band of the system and other observed members such as 4535 Å (3,15) and 4771 Å (5,17) appear weak. Multiplet no. 5 of *NII* ($3s^3P^0$–$3s^3P$) in the 4601–4631 Å range is very weak, if observable.

Table 2. Calculated values at 8 Å resolution and at various assumed excitation temperatures of vibration and rotation. Relative peak intensities $I(1,3)/I(0,2)$ and $I(2,4)/I(0,2)$ include the overlap of bands; bandwidths (Å) refer to intensities of two-thirds peak value

T_v (°K)	T_r (°K)	$\frac{I(1,3)}{I(0,2)}$	$\frac{I(2,4)}{I(0,2)}$	Band width at two-thirds peak (0,2)	(1,3)	(2,4)
600	600	0·02	—	7	9	—
1200	1200	0·26	0·03	7	9	9
2400	1200	0·91	0·19	6	8	9
2400	2400	0·93	0·36	6	17	22
4800	1200	1·78	0·67	7	7	9
4800	2400	1·74	0·99	7	16	20
4800	4800	1·76	1·17	6	24	36

CHARACTERIZATION OF EXPERIMENTAL DATA

To characterize the excitation observed, one calculates the intensity ratios expected from an artificial model with assumed vibration and rotation excitation temperatures. This choice of comparison does not imply that these conditions existed; but there is available no detailed treatment of vibration and rotation in charge exchange excitation, and the atmospheric temperatures which would be required in calculations of electron excitation are unrealistically high. Synthetic spectra were computed according to the standard equation

$$I(v',v'') = b\nu^4(v',v'')p(v',v'')N_v(v',T_v)N_r(J',T_r).$$

The constant b is a scale factor, and the frequency is approximated by the frequency ν_0 of the band origin. Vibrational transition probabilities $p(v',v'')$ are taken from the calculations of WALLACE and NICHOLLS (1955) and NICHOLLS (1956), which include variation of the electronic transition moment. The vibration and rotation population terms N_v and N_r are the normalized Boltzmann factors pertaining to the excited levels at excitation temperatures T_v and T_r. Intensity alternation is omitted. Constants and methods were taken following HERZBERG (1950).

The temperatures 600, 1200, 2400 and 4800°K were used in combinations of $T_r \leqq T_v$, and the effects of overlapping bands were evident in the plotted synthetic spectra. Because of the varying spacing of rotation structure, these spectra incorporated the modifications of observable intensity distribution caused by an instrumental resolution of 8 Å. Vibration enhancement at high T_v is produced

Table 3. Observed values at 8 Å resolution and indicated 150° WM times. Intensities are averaged over 45° sky segments of the magnetic meridian plane, labelled N, NZ, SZ, S. The second decimal figure is doubtful. Peak intensities I include the overlap of bands; integrated intensities $\bar{I}$ do not. Typical comparative data are given for a normal intense aurora of 17 February

Spec. no.	150° WMT	Sky seg.	$I(0,2)$ (arb. units)	$\frac{I(1,3)}{I(0,2)}$	$\frac{I(2,4)}{I(0,2)}$	Band width at two-thirds peak (0,2)	(1,3)	(2,4)	$\frac{\bar{I}(1,3)}{\bar{I}(0,2)}$	$\frac{\bar{I}(2,4)}{\bar{I}(0,2)}$
6	1757–1807	N	14	1·26	0·74	11	12	15	1·5	0·7
		NZ	17	1·29	0·89	11	18	16	1·3	0·6
		SZ	11	1·73	0·86	12	17	16	1·1	0·5
		S	10	2·1	0·94	14	18	18	1·4	0·5
7	1807–1909	N	23	1·39	0·78	18	17	20	0·9	0·5
		NZ	37	1·57	0·97	18	19	19	0·8	0·5
		SZ	56	1·30	0·91	18	22	17	0·7	0·5
		S	54	1·67	0·89	19	20	20	0·8	0·4
8	1909–1929	N	—	—	—	—	—	—	—	—
		NZ	4	1·25	0·87	—	—	—	—	—
		SZ	10	2·00	0·79	23	12	28	0·9	0·5
		S	21	1·76	0·71	10	14	20	1·0	0·6
13	2329–0109	N	78	0·88	0·34	13	17	21	0·5	0·1
		NZ	88	0·73	0·24	12	19	24	0·6	0·2
		SZ	140	0·68	0·21	11	17	26	0·5	0·2
		S	170	0·88	0·34	18	18	21	0·6	0·2
14	0109–0129	N	8	0·94	0·35	11	14	—	0·8	0·3
		NZ	6	1·27	0·54	22	17	—	1·1	0·4
		SZ	16	0·77	0·26	9	14	—	0·7	0·3
		S	22	1·14	0·41	12	14	12	0·8	0·3
(Normal intense aurora 17/3/58)										
011		NZ	66	0·42	0·09	13	12	—	0·4	0·09
		S	84	0·46	0·07	10	12	—	0·5	0·09

largely by the N_v factor in the case of $I(1,3)/I(0,2)$; the next ratio $I(2,4)/I(0,2)$ is largely influenced by N_r at high T_r, for there are considerable contributions from the neighbouring highly developed rotation bands. Table 2 summarizes the calculated ratios of peak intensities at various excitation temperatures, including effects of overlapping and of medium resolution. No sunlit conditions are assumed.

VOL. 16 1959

It is quite likely that in this aurora, because of the importance of charge exchange collisions, the excited levels are actually not populated according to any Boltzmann distribution.

Table 3 gives the measured peak intensity ratios obtained from five exposures at the indicated times and elevations. The first three spectra (nos. 6, 7, 8) represent early evening and partially sunlit conditions, and the other two (nos. 13, 14) are definitely not sunlit. Average values of the ratios would encompass too many variations to be very significant, but they may be estimated as $I(1,3)/I(0,2) = 1{\cdot}3$ (av.) and $I(2,4)/I(0,2) = 0{\cdot}6$ (av.). These spectra are typical of the strong early and middle stages of the display; later morning spectra are weaker but show some persistence of vibration enhancement.

For comparison, the intensity ratios for spectra from a normal intense aurora on 17 February are included in the table and are seen to be appreciably lower.

Ratios of integrated intensities $\bar{I}$ are also listed. These were obtained as areas of intensity profiles with exclusion of estimated overlap contributions.

Comparative tabulation according to observed peak intensities is preferable to the use of integrated intensities because of less involvement with the highly extensive and overlapping rotation structure. This high rotation can account for the general rise in background; it occurs abruptly at the (0,2) band head, is not due to the sunlight continuum, and is also seen in the synthetic spectra at high rotation. While the instrument is capable of resolving the P- and R-branches, they can scarcely be distinguished in the observed spectrograms, a fact which could follow from a non-Boltzmann distribution of rotation.

An approximate comparison of band widths can be made between the synthetic and observed spectra in order to give an indication of rotation enhancement. The order of accuracy cannot be high at 8 Å resolution. To reduce effects of general background, total widths were measured at the level of two-thirds peak intensity. Figs. 6 and 7 show the observed intensity ratios plotted against the width ratios for (1,3) and (2,4) relative to (0,2). Comparative points from the 17 February spectrograms are also shown. The corresponding values obtained from calculated spectra are shown by lines. The peak ratio $I(1,3)/I(0,2)$ is seen to be chiefly dependent on vibration and little affected by rotation, while the peak ratio $I(2,4)/I(0,2)$ shows stronger interaction. One may roughly characterize the varying enhancement of vibration and rotation as corresponding to Boltzmann distributions at excitation temperatures of (1) $T_v = 4000°\mathrm{K}$ and $T_r = 2500°\mathrm{K}$ during the early stages of the display when the aurora was sunlit, purely of type-A, and the hydrogen emissions were intense and (2) $T_v = 2400°\mathrm{K}$, $T_r = 2000°\mathrm{K}$ in later stages when the aurora was a mixture of type-A and type-B, and the hydrogen emissions were faint. In comparison, the normal spectrum of 17 February may be characterized by excitation temperatures of $T_v = 1500°\mathrm{K}$ and $T_r = 1500°\mathrm{K}$.

Although peak intensities have been used in the data of Table 3, it is possible to calculate relative rates of population of the vibration levels. The standard equation as given by Bates (1949) applies,

$$\bar{I}(v',v'') = \frac{cg(v')p(v',v'')\nu^4(v',v'')}{\Sigma_{v''}p(v',v'')\nu^3(v',v'')}.$$

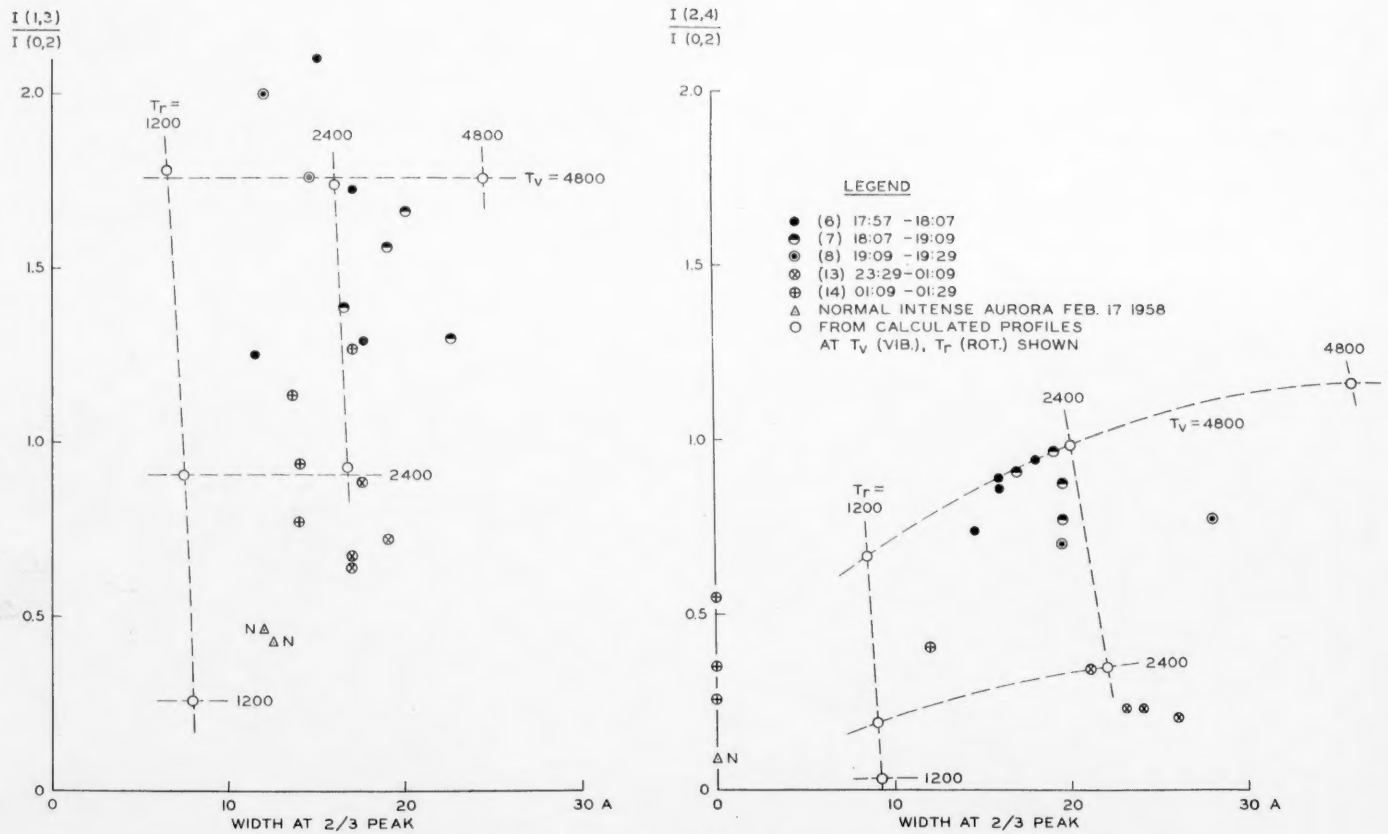


Fig. 6, 7. Enhanced ratios of peak intensities $I(1,3)/I(0,2)$ and $I(2,4)/I(0,2)$ at 8 Å resolution vs. band widths at two-thirds peak. Points show observed values in quarter sections of sky at indicated times. Non-measurable widths are plotted as zero. Lines show values from synthetic profiles assuming Boltzmann excitation temperatures T(vib.) and T(rot.).

VOL. 16 1959

Here I is the vibration band intensity integrated over rotation structure, g the rate of population and c a scale factor. If the transition probabilities p are previously known, it is unnecessary to evaluate them by intensity measurements of all transitions from a common v' level. There is, however, some disagreement in the calculated probabilities at high vibration levels as given by previous authors, and the present report refrains from emphasizing tabulations of derived $g(v')$ values. For equal rates of population of all v' levels, the integrated intensities for

Table 4. Reported relative intensities of the $\Delta v = 2$ sequence

Reference	Type	Relative intensities
Bates, 1949	calc. 4000°, electron excit.	100, 87, 64
Bates, 1949	calc. 8000°, electron excit.	100, 127, 136
Bates, 1949	calc. sunlit	100, 74, 78
Störmer, 1939	obs. blue sunlit, Tromso	100, 97, 85
Rayleigh, 1922	obs. S. England	100, 150, 100
Petrie and Small, 1953	obs. Saskatoon	100, 33
Vegard, 1940	obs. av. "normal"	100, 59, 44
Vegard, 1958	obs. Tromso	100, 64, 45
Barbier and Williams, 1950	obs. College, estimates	9, 7, 2
Clark and Belon (This paper)	obs. College	100, 130, 60
Smyth and Arnott, 1930	lab. 6 keV can. rays, "densities"	7, 7, 9
Fan, 1956a	lab. 20 keV H^+ in air	100, 62, 34
Fan, 1956a	lab. 150 keV He^+ in air	100, 100, 50

the (0,2), (1,3) and (2,4) bands would be predicted as 1·0, 3·6, 2·4 following Wallace and Nicholls (1955) and Nicholls (1956); 1·0, 1·4 following Petrie and Small (1953) using values from Turner and Nicholls (1951); and 1·0, 1·9, 2·9 according to Bates (1949).

A summary of published observed and calculated intensity ratios is given in Table 4. Disparities due to unknown details of dispersion, resolution, profile integration and density calibration are probably less than the actual variations in reported spectra.

Discussion

Auroral emission of the first negative bands of nitrogen, $N_2^+(B^2\Sigma_u^+ \rightarrow X^2\Sigma_g^+)$, is now generally attributed (Meinel, 1952) to a mixture of excitation processes, whose relative importance is subject to considerable variation. Collisions with energetic secondary electrons, which can easily occur in aurorae, lead to ordinary excitation by the process

$$e + N_2(X) \rightarrow N_2^+(B) + 2e. \qquad (1)$$

Because the small electron mass does not affect molecular vibration, the relative intensities of vibration bands can be calculated (Nicholls, 1956; Bates, 1949) using computed transition probabilities and assumed temperatures. Recent photoelectric measurements by Hunten (1955) agree well with this process.

Enhancement of intensity by sunlight leads to considerable vibration. The process

$$h\nu + N_2^+(X) \rightarrow N_2^+(B) \tag{2}$$

is favoured over excitation from the neutral molecule by OLDENBERG (1952) and by BATES (1949), who calculates relative vibration intensities expected at various temperatures. He shows that the intensity measurements of high sunlit rays by STÖRMER (1939) are consistent with theory if one allows for the new population equilibrium in sunlight.

Ionization by high speed protons or other positive ions can be expected (MASSEY and BURHOP, 1952) to operate similarly to reaction (1) by the process

$$H^+ + N_2(X) \rightarrow N_2^+(B) + H^+ + e. \tag{3}$$

SEATON (1953) feels that a considerable proportion of the observed intensity can be attributed to such proton impact ionization. Measurements by FAN (1956a) demonstrate the similarity in excitation by ions and electrons of equal high speeds.

The well observable Doppler shift of hydrogen emission in aurorae (MEINEL, 1951) has led in recent years to widespread consideration of excitation by the charge exchange reaction

$$H^+ + N_2(X) \rightarrow N_2^+(B) + H. \tag{4}$$

General details of such reactions have been described by MASSEY and BURHOP (1952), BATES and DALGARNO (1956), and FAN (1956a,b). Total spin is conserved; and since the net internal energy change is shared as kinetic energy among the three bodies in such a way that momentum is conserved, considerable enhancement of vibration and rotation is to be expected. Applications to auroral excitation are given by GHOSH (1956), MEINEL (1952), PETRIE (1951), SHKLOVSKY (1951) and KRASSOVSKY (1957). Reaction (4) is likely at low proton energies. STIER and BARNETT (1954, 1956) have observed that the charge exchange cross-section for protons in nitrogen reaches a maximum around 1000 km/sec which is somewhat larger than gas-kinetic. MEINEL's (1951) observation on H_α showed speeds up to 3000 km/sec, and those reported in the present paper extend to at least 800 km/sec. STIER and BARNETT (1954, 1956) also observe that for velocities below 3000 km/sec the hydrogen atom is more often neutral than ionized. CHAMBERLAIN (1954, 1958) estimates that a proton will undergo about 700 capture and loss events before stopping, and since the energy change in reaction (4) is small only for capture into the ground state, it is probable that the observed Balmer emission occurs by collision excitation of the incident neutral atom. It is not frequently emphasized that, when neutral, the atom may be effective in exciting forbidden auroral lines through spin exchange collisions.

In the spectra presented in this paper it appears that a strong hydrogen bombardment is closely associated with the auroral features described. As reference to Figs. 2 and 3 will show, there is strong vibration development in the N_2^+ first negative bands when strong H_α and H_β are present. Simultaneous correlation cannot be established in detail, however, and the hydrogen lines seem to weaken in the later hours somewhat more rapidly than does the enhancement. The spectra of enhanced vibration resemble those given by FAN (1956a,b) for laboratory

VOL. 16 1959

protons of 2000 km/sec velocity. From his discussion of energy transferred in the charge exchange collision of two atoms, one can compute that in typical encounters of protons and N_2 there can reasonably be enough energy shared between molecular motions to excite the first few vibration levels and comparable (high) rotation levels. The range of velocities of protons penetrating the atmosphere is sufficient to encompass optimum conditions for such enhancement, which presumably were not exactly met in the above laboratory experiments. Although in the laboratory a greater enhancement was produced using the heavier He^+ ions, there is no reason to construe the present auroral data as indicating the presence of helium. Aside from these observations which would support the concept of charge exchange excitations, it is interesting to consider that the abundant incident hydrogen in its neutral state might, through spin exchange collisions, excite the enhanced atomic states of forbidden multiplicities.

Conclusion

The unusual enhancement of vibration observed in the aurora of 10 February 1958 can plausibly be attributed to the increased importance of charge exchange during the strong proton shower. Sunlight illumination may have contributed to the observed vibration in early evening but cannot have been effective at most of the night-time hours and elevation angles.

It is quite clear from the various published intensity ratios that the predominant excitation mechanism for the auroral N_2^+ bands is subject to change. In view of the variations observable at College, it seems evident that any secure inferences as to latitude effects should be based on statistics. The time averaging in photographic exposures, as Meinel (1952) points out, can easily fail to show short intervals of differing excitation. The anomalous vibration may be more common than supposed but because of short duration is generally not observed photographically except in storms of unusual duration and strength. Indeed, a study of records made in March and April at College with a Hunten photoelectric rapid scanning spectrometer showed several short periods of this same enhancement in auroral arcs. For from ten to fifteen successive 10 sec scans the intensities became reversed and then resumed normal relative values. The intensity of H_β at the same elevation angle was strong but variable. It is planned to study these features systematically in the coming season.

References

Barbier D.	1947	*Ann. Geophys.* **3,** 277.
Barbier D. and Williams D. R.	1950	*J. Geophys. Res.* **55,** 401.
Bates D. R.	1949	*Proc. Roy. Soc.* A **196,** 217, 562.
Bates D. R.	1954	*The Earth as a Planet* (Edited by Kuiper G. P.) p. 576. University of Chicago Press, Chicago.
Bates D. R. and Dalgarno A.	1956	*The Airglow and the Aurorae* (Edited by Armstrong E. B. and Dalgarno A.) p. 328. Pergamon Press, London.
Belon A. E. and Clark K. C.	1959	*J. Atmosph. Terr. Phys.* **16,** 220.
Branscomb L. M., Shalek R. J. and Bonner T. W.	1954	*Trans. Amer. Geophys. Un.* **35,** 107.
Carleton N. P.	1957	*Phys. Rev.* **107,** 110.

Chamberlain J. W.	1954	*Astrophys. J.* **120,** 360.
Chamberlain J. W.	1958	*Advances in Geophysics* (Edited by Landsberg H. E. and Van Mieghem J.) Vol. IV, p. 109. Academic Press, New York.
Chamberlain J. W. and Meinel A. B.	1954	*The Earth as a Planet* (Edited by Kuiper G. P.) p. 514. University of Chicago Press, Chicago.
Dodson H. W., Hedeman E. R. and Owren L.	1953	*Astrophys. J.* **118,** 169.
Fan C. Y.	1956a	*Phys. Rev.* **103,** 1740.
Fan C. Y.	1956b	*The Airglow and the Aurorae* (Edited by Armstrong E. B. and Dalgarno A.) p. 276. Pergamon Press, London.
Ghosh S. N.	1956	*J. Geophys. Res.* **61,** 193.
Herzberg G.	1950	*Spectra of Diatomic Molecules* (2nd. Ed.). Van Nostrand, New York.
Hunten D. M.	1955	*J. Atmosph. Terr. Phys.* **7,** 141.
Hunten D. M.	1956	Fin. Rep. Contract AF 19(122)-152 p. 26. University of Saskatchewan, Saskatoon.
Krassovsky V. I.	1957	Paper presented at Toronto meeting of Internat. Assoc. of Geomag. and Aeronomy, Sept. 1957.
Massey H. S. W. and Burhop E. H. S.	1952	*Electronic and Ionic Impact Phenomena.* Clarendon Press, Oxford.
Meinel A. B.	1951	*Astrophys. J.* **113,** 50.
Meinel A. B.	1952	*Mém. Soc. Sci. Liège* **12,** 203.
Montalbetti R.	1957	*Canad. J. Phys.* **35,** 831.
Nicholls R. W.	1956	*The Airglow and the Aurorae* (Edited by Armstrong E. B. and Dalgarno A.) p. 302. Pergamon Press, London.
Oldenberg O.	1952	*Z. Phys.* **133,** 15.
Petrie W.	1951	Sci. Reps. AR-1 and AR-4, University of Saskatchewan, Saskatoon.
Petrie W. and Small R.	1952	*J. Geophys. Res.* **57,** 51.
Petrie W. and Small R.	1953	*Canad. J. Phys.* **31,** 911.
Petrie W.	1953	*J. Atmosph. Terr. Phys.* **4,** 5.
Lord Rayleigh	1922	*Proc. Roy. Soc.* A **101,** 114, 312.
Roesler F. L., Fan C. Y. and Chamberlain J. W.	1958	*J. Atmosph. Terr. Phys.* **12,** 200.
Romick G. J.	1959	Catalogue of time-resolved zenith auroral spectra, College, 1957–58. Geophys. Institute, College, Alaska. In preparation.
Seaton M. J.	1953	*J. Atmosph. Terr. Phys.* **3,** 285, 295.
Seaton M. J.	1956	*The Airglow and the Aurorae* (Edited by Armstrong E. B. and Dalgarno A.) p. 228. Pergamon Press, London.
Shepherd C. G. and Hunten D. M.	1955	*J. Atmosph. Terr. Phys.* **6,** 328.
Shklovsky I. S.	1951	*Dokl. Acad. Nauk SSSR* **81,** 4, 525.
Smyth H. D. and Arnott E. G. F.	1930	*Phys. Rev.* **36,** 1021.
Stier P. M., Barnett C. F. and Evans G. E.	1954	*Phys. Rev.* **96,** 973.
Stier P. M. and Barnett C. F.	1956	*Phys. Rev.* **103,** 896.
Störmer C.	1939	*Terr. Mag. Atmos. Electr.* **44,** 7.

VOL.
16
1959

Turner R. G. and Nicholls R. W.	1951	*Phys. Rev.* **82,** 290.
Vallance Jones A. and Harrison A. W.	1955	*J. Atmosph. Terr. Phys.* **6,** 336.
Vegard L. and Kvifte G.	1945	*Geofys. Publ.* **16,** No. 7.
Vegard L.	1940	*Terr. Mag.* **45,** 5.
Vegard L.	1958	*Geofys. Publ.* **20,** No. 4.
Wallace L. V. and Nicholls R. W.	1955	*J. Atmosph. Terr. Phys.* **7,** 101.
Warwick C. A.	1954	*Astrophys. J.* **120,** 237.

L.
6
59

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 220 to 227. Pergamon Press Ltd. Printed in Northern Ireland

Spectroscopic observations of the great aurora of 10 February 1958—II

Unusual atomic features

A. E. Belon and K. C. Clark*
Geophysical Institute, University of Alaska, College, Alaska

(*Received* 16 *December* 1958)

Abstract—Auroral spectrograms taken with the IGY patrol spectrograph at College, Alaska, during the great red aurora of 10–11 February 1958, show a large number of seldom seen atomic lines. Most of these enhanced emissions are identified as transitions of O*I* and N*II*. In particular, the identification of both low level forbidden lines of N*II* at 5755 Å and 6584 Å definitely establishes their presence in the auroral spectrum. Relative intensities of the auroral atomic emissions in three representative spectra are given and are compared with their intensities in a normal intense aurora. The unusual intensities and spatial behaviour of some emissions are discussed in terms of their excitation mechanisms.

Introduction

During the night of 10–11 February 1958, a great red aurora was observed in the western hemisphere as far south as Cuba and Mexico City, and as far north as Ice Station A (latitude 80° N). At College, Alaska, the aurora was first observed soon after sunset. In addition to its great brightness, the main visual characteristics were, during the first few hours, its exclusively red colour, its large extent, which at times was up to 95 per cent of the sky, its extreme diffuseness, and its apparent lack of motion.

VOL. 16 1959

Spectroscopic observations of this aurora at College are the subject of two publications. The present paper is concerned with the identification and behaviour of a large number of enhanced atomic emissions seldom seen in intense auroral spectrograms. Part I of this paper (Clark and Belon, 1959) discusses the observed high vibrational excitation in the $B\,(^2\Sigma_u^+)$ state of N_2^+ as indicated by the enhanced intensities of the 4652 Å (1,3) and 4600 Å (2,4) bands relative to the 4709 Å (0,2) band of the first negative system.

Eighteen successive grating spectrograms were obtained throughout the night from 1708 to 0529 hours 150° WMT† with the IGY patrol spectrograph. The entire sequence of these spectrograms together with properly spaced all-sky photographs and pertinent visual observations is presented in Fig. 2 of Part I. In the present paper three of these spectra are selected for principal study. These spectra, illustrated in Fig. 1, are typical of the three phases of the auroral development, namely, the twilight period, the high altitude sunlit period and the active non-sunlit period. The entire sequence of spectrograms, summarized in a later table, is used to support the reliability of reported lines in the above typical spectra.

Methods and Analysis

The IGY patrol spectrograph has an effective aperture of $f/0{\cdot}625$ and covers the spectral region from 3500 Å to 6800 Å with an average dispersion of 330 Å/mm

* On leave from University of Washington, Seattle, 1957–1958.
† 150° West Meridian Time is used here throughout.

and a resolution of 8·2 Å. An all-sky field lens places on the entrance slit an image of the sky extending from the north horizon to the south horizon along the geomagnetic meridian with fiducial marks at 45° elevations. For convenience of analysis, each of the eighteen spectrograms is considered as consisting of four separate but simultaneous spectra, two covering the northern half of the sky and the other two the southern half. These will be denoted as north (N), north zenith (NZ), south zenith (SZ) and south (S), giving a total of seventy-two spectrograms.

The spectrograph produces after each sky exposure the spectrum of a tungsten lamp reduced through a stepped neutral filter. In addition, some exposures were obtained using a standard lamp of known temperature. By means of these various calibration exposures it was possible to correct photographic densities for the characteristic curve of the emulsion, the spectral sensitivity of the 103a-F film, and radiation losses in the spectrograph. All auroral spectrograms were analyzed with a Leeds and Northrup recording microdensitometer. Fig. 1 shows three representative microdensitometer tracings along with photographs of their respective spectra. A plastic template calibrated according to wavelength and overall sensitivity was used on the tracings to convert line densities to relative intensities. Because of the medium resolution of the spectrograph only peak intensities were measured, and no attempt was made to integrate the areas under the line profiles.

Results

The atomic lines present in the spectrograms are listed in Table 1. In determining the elements to which these lines are attributed and the excitation potential of the upper energy levels involved, extensive use was made of lists of auroral wavelengths by Chamberlain and Oliver (1953), Vallance Jones (1955), Vegard and Kvifte (1951) and Petrie and Small (1953). Table 1 also lists the total number of patrol spectrograms for this night in which a given atomic line was observed. In addition, the atomic line intensities in the three spectra illustrated in Fig. 1 are given and are compared with the intensities observed in the spectrum of a normal bright aurora taken on 17 February 1958. The mercury lines present on the spectra are a contamination due to distant street lights.

Table 2 is a graphical representation of the time and position of the atomic emissions observed in the eighteen multiple spectrograms. It must be pointed out that even if a given emission is in the auroral spectrum, its presence or absence in the recorded spectra partly depends on the duration of the exposure, the presence of obscuring features, and the twilight background. This is an inherent limitation in the figures and tables presented.

O*I lines*

The most intense features are the usual lines at 5577 Å, 6300 Å and 6364 Å. In every one of the spectrograms the nebular red doublet is greatly enhanced and more intense than 5577 Å. The first spectrogram, exposed from 1707 to 1708 150° WMT, shows the largest enhancement with an intensity ratio $I_{6300,6364}/I_{5577} = 6{\cdot}7$.

In addition to these three low level forbidden emissions, most of the strongly exposed spectrograms show the presence of permitted transitions from both the

Table 1. Classification of atomic lines observed in the seventy-two auroral spectrograms. Relative intensities for three typical periods are given with no correction for duration of exposure. For comparison the lines observed in an ordinary intense aurora (17 February) are also listed

Elem.	Transition	Exc. pot. (eV)	Reference*	λ	No. of spectra where present	Observed intensities			
						Spectrum 3 1729–1747 NZ 18 min	Spectrum 7 1807–1909 S 62 min	Spectrum 13 2329–0109 S 100 min	Spectrum normal intense 100 min
OI	$2p^4\ ^3P$–$2p^4\ ^1D$	1·96	2, 9, 11, 5, 7	6300	72	overexposed	overexposed	overexposed	overexposed
			2, 9, 11, 5, 7	6364	72	overexposed	overexposed	overexposed	overexposed
	$2p^4\ ^1D$–$2p^4\ ^1S$	4·17	2, 9, 11, 5, 7	5577	72	overexposed	overexposed	overexposed	overexposed
	$3p\ ^5P$–$5s\ ^5S^0$	12·61	2, 11, 5, 7	6455	37	4	12	20	—
	$3p\ ^5P$–$4d\ ^5D^0$	12·70	2, 11, 5, 7	6157	32	3	10	10	1
	$3s\ ^3S^0$–$5p\ ^3P$	12·82	2, 11, 5, 7	3692	10	trace	trace	trace	—
	$3p\ ^5P$–$6s\ ^5S^0$	12·96	2, 11, 5, 7	5436	11	—	4	4	1
	$3p\ ^3P$–$6s\ ^3S^0$	12·98	2, 11, 5, 7	6046	17	—	3	10	2
	$3p\ ^5P$–$5d\ ^5D^0$	13·01	2, 11, 5, 7	5330	8	—	4	4	trace
	$3p\ ^5P$–$6d\ ^5D^0$	13·18	2, 11, 5, 7	4968	7	—	1	1	—
OII	$2p^3\ ^4S^0$–$2p^3\ ^2D^0$	3·31	2, 11, 5	3727	17	trace	weak	weak	—
	$3s\ ^4P$–$3p\ ^4D^0$	25·54	2, 11, 5	4676	40	trace	5	4	trace
	$3s\ ^4P$–$3p\ ^4P^0$	25·74	2, 11, 5, 7	4320	55	2	3	5	1
	$3s\ ^2P$–$3p\ ^2D^0$	26·14	2, 11, 5, 7	4416	47	8	28	30	2
	$3p\ ^4P^0$–$3d\ ^4D$	28·73	2, 11, 5	4120	18	—	trace	trace	trace
	$3s\ ^6S$–$3p\ ^6P$	33·19	9, 11	4465	34	trace	7	7	—
NI	$2p^3\ ^4S^0$–$2p^3\ ^2D^0$	2·37	2, 11, 5, 7	5200	57	22	13	54	6
	$3s\ ^4P$–$4p\ ^4S^0$	13·26	2, 11, 5, 7	4151	23	—	2	1	—
NII	$2p^2\ ^3P$–$2p^2\ ^1D$	1·89		6584	18	29	obscured	obscured	—
	$2p^2\ ^1D$–$2p^2\ ^1S$	4·04	2, 11, 6	5755	20	4	1	5	—
	$3s\ ^1P^0$–$3p\ ^1P$	20·32	2, 5, 7	6482	27	4	10	10	—
	$2p^3\ ^1D^0$–$3p\ ^1P$	20·32	2, 11, 5, 7	4895	5	—	2	2	—
	$3s\ ^3P^0$–$3p\ ^3D$	20·57	2, 11, 5	5710	10	—	3	4	1
			2, 11, 5	5676	49	8	30	53	4
			2, 11, 5, 7	5666	27	—	trace	trace	trace
	$3p\ ^3D$–$3d\ ^3F^0$	23·04	2, 11, 5, 7	5001	51	16	41	64	4
	$3p\ ^3P$–$3d\ ^3D^0$	23·14	2, 11, 5	5961	19	3	trace	5	4
	$3p\ ^3D$–$3d\ ^3D^0$	23·14	2, 11, 5	4780	6	—	—	2	—
			2, 11, 5	4791	2	—	—	1	—
	$3p\ ^3D$–$3d\ ^3P^0$	23·31	2, 11, 5	4507	24	10	18	7	—
			2, 11, 5	4488	21	—	2	4	—
	$3d\ ^1F$–$4f\ ^1G$	26·1	9, 11	4535	37	—	11	12	1
	$3d\ ^3P$–$4f\ ^3D$	26·1	9, 11	4432	39	trace	11	15	—
	$3s\ ^5P$–$3p\ ^5D$	27·7	11	5535	9	—	5	3	—
H	H_α $2\ ^2P^0$–$3\ ^2D$	12·04	2, 11, 5, 7	6563	29	40	50	obscured	absent in this particular spectrum
	H_β $2\ ^2P^0$–$4\ ^2D$	12·69	2, 11, 5, 7	4861	23	5	13	4	
Na	$3\ ^2S$–$3\ ^2P^0$	2·10	2, 11, 5	5890	40	2	4	10	18
Unidentified features	—	—	—	3780	26	strong	—	—	—
	—	—	—	3812	11	strong	—	—	—

* (2) Chamberlain and Oliver, 1953; (5) Vallance Jones, 1955; (6) Petrie, 1952; (7) Petrie and Small, 1953; (9) Vegard, 1958; (11) Vegard and Kvifte, 1951.

VOL. 16 1959

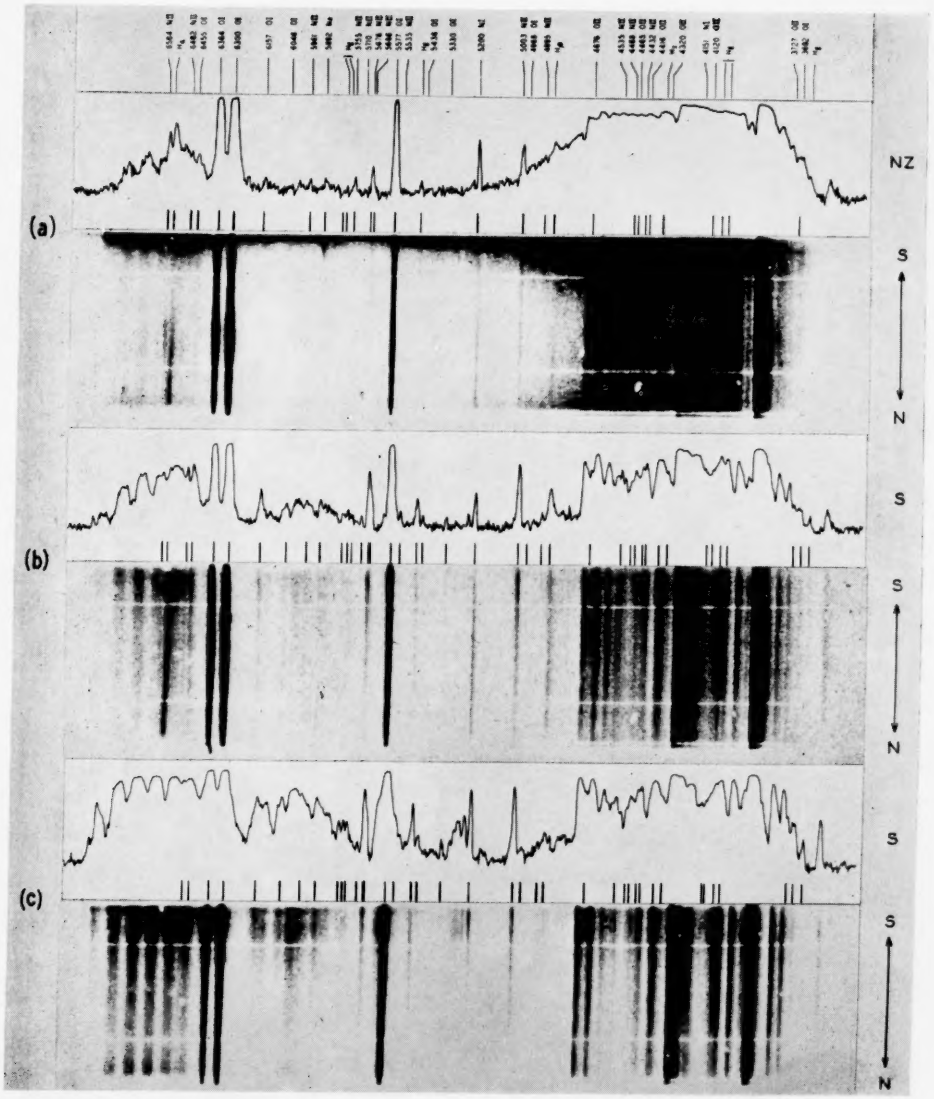


Fig. 1. Spectrograms used in primary analyses. Elevations in the magnetic meridian, principal atomic lines, and typical densitometer traces are shown. (a) Spectrum 3, 1729–1747; (b) spectrum 7, 1807–1909: (c) spectrum 13, 2329–0109.

VOL.
16
1959

Table 2. Time and position of occurrence of the various atomic emissions observed. In each spectrum the magnetic meridian has been divided into four equal parts identified from left to right as N, NZ, SZ, S. Definite presence of the emission is indicated by an "x". The "–" indicates that the emission is not detected

Spectrum no.		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Exposure start (150° WMT)		1708	1709	1729	1747	1752	1757	1807	1909	1929	2109	2129	2309	2329	0109	0129	0309	0329	0509
Exposure duration (min)		01	20	18	05	05	10	62	20	100	20	100	20	100	20	100	20	100	20
OI	6300	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx
	6364	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx
	5577	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx
	6455	–	–	xxxx	–	–	–	xxxx	–xx	–xx	–xx	xxxx	–	xxxx	xxxx	–xxx	xxxx	xxxx	–
	6157	–	xx–	xxxx	–	–	–	xxxx	–xx	–xxx	–xx	xxxx	–	–xxx	–	–xx	xxxx	–xx	–
	3692	–	–	–	–	–	–	xxxx	–	–xx	–	–	–	–xxx	–	–x	–	–	–
	5436	–	–	–	–	–	–	–xx	–	–xx	–	–xx	–	–xxx	–	–x	–	–x	–
	6046	–	–	–	–	–	–	xxxx	–	–xx	–	–xxx	–	xxxx	–	–xx	–x	–x	–
	5330	–	–	–	–	–	–	–xx	–	–xx	–	–x	–	–x	–	–x	–	–x	–
	4968	–	–	–	–	–	–	–xx	–	–xx	–	–	–	–x	–	–x	–	–x	–
OII	3727	–	–	xxx–	–	–	–	–xxx	–	x–xx	–	–xx	–	xxxx	–	–xx	–	–	–
	4676	–	–	xxx–	xxxx	xxxx	xxxx	xxxx	–xxx	x–xx	–xx	–xx	–	xxxx	–xx	–x–	–	–xx	–xx
	4320	–	–	xxx–	xxxx	–xxx	xxxx	xxxx	–xxx	xxxx	xxxx	xxxx	–x	xxxx	xxxx	xxxx	xxxx	xxxx	–x
	4416	–	–	xxx–	xxxx	–xx–	xxxx	xxxx	–xx	–xx	–xx	–xx	–x	xxxx	xxxx	xxxx	xxxx	xxxx	–x
	4120	–	–	–x–	–xx	–	–x	–xx	–xx	–x	–xx	–	–	–x–x	–x	–xx	–x–x	–	–
	4465	–	–	xxx–	–	–x–	–xxx	xxxx	–xx	–xxx	–x–x	–xxx	–	–xxx	–xx	–xxx	–xxx	–xx	–
NI	5200	–	xxx–	xxxx	xxx–	xx–	xxxx	xxxx	–xxx	xxxx	xxxx	xxxx	–x	xxxx	–xxx	xxxx	xxxx	xxxx	–xx
	4151	–	–	–	–xxx	–xxx	xxxx	xxxx	–x	x–xx	–x	–	–	xxxx	–	–	–	–	–
NII	6584	–	xxx–	xxxx	–	–	xxx–	xxxx	–	–x–	–	–x–	–	–	–	–xx	–	–	–
	5755	–	xxx–	x––	–	–	–x–	–xxx	–	–xx	–x–	–xx	–	xxxx	–	–xx	–	–x	–
	6482	–	–	xxxx	–	–	–	xxxx	–x	–xx	–xx	–x	–	xxxx	–xx	–xx	–x	xxxx	–
	4895	–	–	–x–	–	–	–	–xx	–	–	–	–	–	–x	–	–	–	–x	–
	5710	–	–	–	–	–	–	xxxx	–	–xx	–	–x	–	–xx	–	–x	–	–	–
	5676	–	xxx–	xxxx	–	–	xxxx	xxxx	–xx	xxxx	–xxx	–xxx	–x	xxxx	xxxx	xxxx	xxxx	xxxx	–x
	5666	–	–	xxxx	–	–	xxxx	xxxx	–x	–xx	–xx	–xxx	–	xxxx	–	–x	–x	–x	–
	5001	–	–	xxxx	xx–	–xx–	xxxx	xxxx	–xxx	xxxx	xxxx	xxxx	–	xxxx	–xxx	xxxx	xxxx	xxxx	–x
	5961	–	–	–x–	–	–	–	xxxx	–	–xx	–xxx	–x	–	xxxx	–	–xx	–	–xx	–
	4780	–	–	–	–	–	–	–	–	–xx	–	–	–	xxxx	–	–	–	–	–
	4791	–	–	–	–	–	–	–	–	x–	–	–	–	–x	–	–	–	–	–
	4507	–	–	xxx–	–	–x–	–xx	xxxx	–	–xx	–xx–	–xx	–	xxxx	–	–xx	–	–xx	–
	4488	–	–	xx–	–	–	–	xxxx	–	–xx	–xxx	–xx	–	xxxx	–	–xx	–	–x–x	–
	4535	–	–	–	–	–x–	–xxx	xxxx	–xx	xxxx	xxxx	–xxx	–	xxxx	–xx	xxxx	–xxx	–xxx	–
	4432	–	–	xxx–	xxxx	–x–	xxxx	xxxx	–xx	–xx	–xx	–xx	–	xxxx	–xx	x–xx	xxxx	–xx	–
	5535	–	–	–	–	–	–	–xx	–	–xx	–	–x	–	–xxx	–	–x	–	–	–
H	6563	x–	xxx–	xxxx	xx–	xxxx	xxxx	xxxx	xxxx	xxx–	–	–	–	–	–	–	–	–	–
	4861	–	–	xxx–	–x–	xxx–	xxxx	xxxx	–xx	–xxx	–	–x	–	–xx	–	–	–	–	–
Na	5890	xxxx	xxx–	xxxx	–	–	–	xxxx	–	xxxx	–xxx	xxxx	–	xxxx	–	xxxx	–xx	xxxx	–
Unident. features	3812	xxxx	xxx–	xxxx	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	3780	xxxx	xxx–	xxxx	–	–	–	–	–	xx–	–x	–xx	–	–	–x	xxxx	x–x	xx–	–x

quintet and triplet levels. All quintet emissions listed are, in most exposures, more intense than any neighbouring spectral features, and their presence is therefore well established. Of the two triplet transitions observed, only the emission at 6046 Å can be reported with certainty; the very faint emission at 3692 Å is usually obscured by the well developed second positive band of N_2 at 3710 Å (2,4) and is observable on only ten densitometer tracings. The intensity ratio of the 5436 Å quintet line relative to the corresponding 6046 Å triplet is variable. This ratio

rises slightly above unity on spectrum 7 associated with intense hydrogen emission and becomes appreciably smaller on later spectra where the hydrogen emission is faint. Since these two permitted lines come from levels of differing multiplicity but of almost the same excitation energy, the variation in relative intensities may indicate a change in the excitation processes. For example, the quintet levels, excitable through recombination but not through direct electron impact, may also be appreciably populated through electron exchange with fast free electrons or with fast neutral atoms in the partly ionized flux of hydrogen.

OII lines

The *OII* lines occur in a region of the visible auroral spectrum which is occupied by a large number of other spectral features. This fact, coupled with the moderate resolution of the spectrograph, makes it difficult to distinguish many of the *OII* lines from the background of other lines and bands. Therefore, Table 1 lists only those *OII* lines which are easily distinguishable, although they may not be the strongest members of a given multiplet.

In contradistinction to the general great enhancement of the nebular transitions of *OI*, *NI*, and *NII* during the aurora, the nebular *OII* doublet at 3727 Å is very weak. Its relative intensity could not be well measured because of calibration difficulties at wavelengths shorter than 4000 Å.

The moderately intense emission at 4465 Å has not been previously detected at College. It has, however, been reported and identified by Vegard and Kvifte (1951) and Vegard (1958) as originating from a sextet level of *OII*.

VOL. 16 1959

NI lines

The only emission of *NI* present with appreciable intensity is the low level nebular transition at 5200 Å. It shows considerable enhancement throughout the night. Its spatial distribution follows closely that of the other nebular lines in the auroral spectrum. The spectra show a striking lack of a consistent spatial relationship between *NI* 5200 Å and the permitted line *NII* 5001 Å. This difference is well illustrated in spectra 7 and 13 of Fig. 1. This observation is, however, difficult to evaluate in terms of excitation processes because of the extremely long lifetime (8 hr) of the $^2D^0$ state of *NI*, which could carry the 5200 Å emission over the durations of several exposures.

A faint feature appearing regularly at about 4150 Å on moderately intense spectrograms has been identified as the $3s\ ^4P$–$4p\ ^4S^0$ transition of *NI*.

NII lines

Most of the 10–11 February spectrograms show a great number of *NII* lines which are usually absent in other intense spectra taken at College. In particular the appearance of the low level forbidden lines at 5755 Å (auroral) and 6584 Å (nebular) is of major importance because it definitely establishes their presence in the auroral spectrum. The auroral transition has been previously observed by Petrie (1952), but he explicitly states that the nebular transition is absent on his spectra. In the present spectra the 6584 Å emission is, in the earlier sunlit exposures, much more intense than the 5755 Å emission. This fact is consistent

with the characteristic enhancement of nebular lines in this aurora and thereby strengthens the identification. The theoretical and observational analogy between the forbidden transitions of N*II* and O*I* is in fact so striking that their excitation mechanisms may have been identical during the early stages of the display. The aurora was then at high altitude and entirely sunlit. Under these conditions the concentration of nitrogen ions may still have been large enough to account for the observed intensities through exchange excitation of the ions. This process would indeed explain the apparent absence of the N*II* forbidden lines in most aurorae, where the operative mechanism would be through a collisional dissociation of the neutral nitrogen molecule (BATES, 1955).

Since, as PETRIE (1952) points out, the excitation of the singlet levels of N*II* could proceed by charge exchange from the metastable doublet levels of N*I* but not from its ground quartet state, it is interesting that the considerable 5200 Å intensity might indicate a sufficient population of metastable atoms for this process. However, if sunlight enhancement of the ion concentration is important, a comparison of spectra 3 and 7 would favour the process of excitation from the ionized state. The sunset shadow lines of Fig. 1, Part I, show an average geometrical shadow height in the North–Zenith of 100–170 km for spectrum 3 and of 180–600 km for spectrum 7. Because of the opacity of the atmosphere to the ionizing wavelengths (DITCHBURN, 1954; CLARK, 1952), the effective shadow lines are considerably higher. Since incident hydrogen and metastable N*I* are present strongly in both spectra, it is the temporary ionization by sunlight which appears to be more closely connected with the great enhancement of the nebular N*II* line in spectrum 3. Nevertheless, the similar early enhancement of the nebular O*I* lines, which in this aurora was much larger than the expected sunlight enhancement, would suggest that other variables, such as changes in height or in energy distribution of incoming hydrogen, may have been effective.

The analogy between N*II* and O*I* does not carry over to the permitted levels. In the case of N*II*, the spectra of 10–11 February, show a large number of transitions from the singlet, triplet and quintet levels. The triplet and singlet emissions are the most prominent. The quintet emission is weak, but very sharp on the spectrograms where it appears. Three of the N*II* transitions listed in Table 1 have been reported and identified previously only by VEGARD and KVIFTE (1951) and VEGARD (1958). Their identification is retained on the basis of the similarity in spatial occurrence of these lines along with other known N*II* lines. It is possible that the O*I* transition $3p\,^3P$–$5d\,^3D$ also contributes to the emission at 5960 Å attributed to N*II* in Table 1.

H *lines*

One of the outstanding characteristics of the aurora of 10–11 February 1958 is the very large influx of low energy hydrogen atoms in the early stages of the display. Doppler shift measurements of H_α are illustrated and discussed in Part I. They indicate a maximum observed velocity of approach of 800 km/sec. Until 1900 150° WMT both H_α and H_β appear intensely even on short exposures. H_γ is customarily obscured by other spectral features of greater intensity and is not listed. After 1900 H_α weakens considerably and is obscured by the increasingly

intense first positive system of N_2. However, H_β remains detectable for most of the night. The marked decrease in the intensity of H lines at the onset of active first positive emission from N_2 has been regularly observed at College.

Unidentified features

Two spectral features at approximately 3812 Å and 3780 Å appear in the spectra which are contaminated by twilight. These are not the band heads of the second positive system of N_2. Their intensity is fairly large as they appear strongly in the first spectrum, taken from 1708 to 1709. They are still visible in the south on spectrum 3 of Fig. 1. These features appear to be of atomic origin, and their spatial distribution shows a strong twilight enhancement.

Discussion

The observations presented in Parts I and II show that the great aurora of 10–11 February 1958 had, in addition to its great brightness, several unusual characteristics. These can be summarized as: (1) the generally high altitude and diffuseness of the display; (2) the presence of hydrogen in great abundance and of possibly slower maximum speed than usual; (3) a general enhancement of the nebular atomic lines; and (4) high vibration in the $B(^2\Sigma_u^+)$ state of N_2^+. At this point it is interesting to try to unify these various observations and speculate on the nature of the conditions which occasionally result in a great aurora such as the one described here.

VOL. 16 1959

In agreement with other studies made in recent years at College (Romick and Elvey, 1958) the hydrogen emissions do not, in general, seem to correlate conclusively in time and position with other emissions. However, there seems to be a close association between the strong hydrogen flux and the general character of this aurora. Part I shows that the enhancement of vibration in N_2^+ can plausibly be attributed to the increased importance of charge exchange between slow incoming protons and the neutral nitrogen molecules over direct ionization and excitation of N_2 by fast protons and electrons. Part I, therefore, tends to connect features (4) and (2) and perhaps also (1) if the low energy of the protons is also an attribute of the primary exciting particle stream.

Collisional de-activation must certainly play a large part in relating features (1) and (3). How large a part it plays can only be answered through future observation in great aurorae of the heights as well as the intensities of the emissions involved. It is doubtful, however, that de-activation is the only consideration, as the presence of unusual lines of other multiplicities and the great intensity of the nebular lines suggest that additional excitation processes may operate. From these spectra one cannot assess the relative importance of the excitation of atomic states by recombination, dissociation, direct impact with electrons or protons, and exchange collisions with electrons or hydrogen atoms. Nevertheless, it is interesting to speculate that the known large flux of low energy hydrogen atoms may have been very effective in exciting the atomic states of forbidden multiplicities through spin exchange collisions. Such a process would correlate features (2), (3) and (1), and would be consistent with (4) because of the many alternations of charge as the proton penetrates the atmosphere.

On the basis of the above discussion it appears possible that such an aurora is caused by an abnormally high influx of protons whose average energy is lower than usual. This will result in a shift in the relative importance of the various auroral excitation mechanisms which can, in principle, give rise to the unusual features observed. A somewhat similar idea has already been advanced by VEGARD (1957). It is planned to investigate this hypothesis during the current season by means of more refined observations and with the help of past spectrograms.

REFERENCES

BATES D. R.	1955	*Ann. Geophys.* **3,** 253.
CHAMBERLAIN J. W. and OLIVER N. J.	1953	*J. Geophys. Res.* **58,** 457.
CLARK K. C.	1952	*Phys. Rev.* **87,** 271.
CLARK K. C. and BELON A. E.	1959	*J. Atmosph. Terr. Phys.* **16,** 205.
DITCHBURN R. W.	1954	*Rocket Exploration of the Upper Atmosphere* (Edited by BOYD R. L. F. and SEATON M. J.) p. 325. Pergamon Press, London.
PETRIE W.	1952	*Phys. Rev.* **87,** 1002.
PETRIE W. and SMALL R.	1953	*Canad. J. Phys.* **31,** 911.
ROMICK G. J. and ELVEY C. T.	1958	*J. Atmosph. Terr. Phys.* **12,** 283.
VALLANCE JONES A.	1955	Scient. Rept. No. AR–20, University of Saskatchewan, Saskatoon
VEGARD L.	1957	*The Threshold of Space* (Edited by ZELIKOFF M.) p. 22, Pergamon Press, London.
VEGARD L.	1958	*Geofys. Publ.* **20,** No. 4.
VEGARD L. and KVIFTE G.	1951	*Geofys. Publ.* **18,** No. 3.

L.
6
59

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 228 to 235. Pergamon Press Ltd. Printed in Northern Ireland

Meteor radiant distributions and the radio-echo rates observed by forward scatter*

M. L. MEEKS and J. C. JAMES
Georgia Institute of Technology, Atlanta, Georgia

(*Received* 26 *November* 1958)

Abstract—The diurnal variations in sporadic meteor echo rate observed by the forward scattering of radio waves over three different paths are compared with predictions based on a simplified model of the average radiant distribution which has been obtained from radar data by HAWKINS. Our simplified distribution is composed of three hypothetical meteor showers with radiants at the apex of the earth's way and at points in the ecliptic at angles $\pm 65°$ from the apex. Reasonably good agreement is found between these predictions and extensive measurements reported by VOGAN and CAMPBELL. Additional measurements reported here for two paths—Boston, Massachusetts to Atlanta, Georgia (1480 km) and Boston, Massachusetts to Columbia, South Carolina (1250 km)—also show similar agreement. Evidence of the concentrations of radiants near the sun and antihelion points appear in the average diurnal rate for periods of about 1 month. In general forward-scatter measurements appear consistent with HAWKINS' radiant distribution.

INTRODUCTION

THE forward scatter of radio waves by meteor trails is of practical interest as a propagation mode for v.h.f. communication. As a consequence of this application a large amount of data has been gathered on meteor-echo rates over various propagation paths. In the past such forward-scatter data have been used by Canadian workers in the study of meteor showers, but a comparatively small amount of effort has been made to interpret the much larger quantity of data on sporadic meteors. It is the purpose of the present paper to examine data on sporadic meteors and to determine the extent to which such data are consistent with present estimates of the average distribution of sporadic meteor radiants.

Forward-scatter data are, in general, much more difficult to interpret than radar data because of purely geometrical complexities. An approximate method for computing the diurnal variations in forward-scatter echo rate for an arbitrary meteor-shower radiant has been given by HINES (1955). This approximation makes use of the simplifying assumption that specular echoes are produced by meteor trails that are tangent to one of a family of cylinders whose axes coincide with the line connecting the transmitter and receiver. The errors introduced by this cylindrical approximation are usually negligible for separations of about 1000 km or more between transmitter and receiver. This approximation has served as a basis for a preliminary investigation of several major streams (FORSYTH *et al.*, 1955) over the 860 km path from Greenwood, Nova Scotia, to Ottawa, Ontario. HINES and VOGAN (1957) have reported detailed observations of the 1956 Quadrantid meteor shower over the same path.

The problem of computing the diurnal variations in forward-scatter echo rate to be expected from an arbitrary meteor shower has been solved exactly by MEEKS and JAMES (1957). Independently and almost simultaneously exact solutions were

* This research was sponsored by the Office of Naval Research, Contract Nonr–991(02), and by the Air Force Cambridge Research Center, Contract AF 19(604)–1593.

VOL.
16
1959

obtained by HINES (1958) for 1000 km paths lying north–south and east–west. With the exact solution it is possible to study meteor showers over paths which are short compared to 1000 km. For such short paths the cylindrical approximation must be replaced by the exact condition for specular reflection: namely, the meteor trail must be tangent to one of a family of spheroids with foci at the transmitter and receiver locations.

Once the problem of computing the echo-rate variations from a single radiant point has been solved, it becomes possible to predict the diurnal variations in echo rate for an arbitrary distribution of sporadic meteor radiants. If a sufficiently large number of point radiants are assumed, the echo-rate variations resulting from a continuous radiant distribution can be predicted with arbitrarily high accuracy. It has therefore seemed desirable to compare the sporadic meteor data, obtained by forward scatter, with simple approximate predictions based on present knowledge of the average sporadic radiant distribution. Fortunately only a small number of point radiants are needed for such an approximation.

Approximate Treatment of the Sporadic Meteor–Radiant Distribution

Radio-echo techniques have made it possible to observe the meteor flux continuously over long periods of time. With these techniques the average distribution of meteor radiants may be investigated over a much larger portion of the celestial sphere than is possible by visual observations. Radar observations of sporadic meteors during the period October 1949 to September 1951 have been analysed in detail by HAWKINS (1956a,b) to obtain statistical estimates of the radiant distribution. A study of 240,000 echoes shows a concentration of radiant points toward the plane of the ecliptic with marked clustering around the apex of the earth's way, near the sun, and near the antihelion point. HAWKINS was able to reconstruct an average radiant distribution from these radar data by assuming the actual distribution to be symmetric with respect both to the ecliptic plane and to a plane perpendicular to the ecliptic and containing the apex of the earth's way. Visual observations reported by HAWKINS and PRENTICE (1957) have shown agreement with the radiant distribution found in this way. The concentration of radiants near the antihelion point is confirmed by the visual observations, and these observations give indications of a somewhat more diffuse concentration of radiants near the apex. The distribution of radiants near the sun of course cannot be studied by visual observations because of the darkness requirement.

It should be emphasized that this picture of the sporadic radiant distribution represents an average taken over a period of more than 1 year, and marked variations might be expected from day to day.

The simplest discrete approximation to HAWKINS' average sporadic radiant distribution consists of three point-radiants located respectively at the apex of the earth's way and on the ecliptic at $\pm 65°$ from the apex. These three points lie exactly at the maxima in the distribution obtained from radar data, and HAWKINS' distribution furthermore indicates that to a first approximation these points should be weighted equally in computing the echo rate. Clearly this is a drastic oversimplification of HAWKINS' distribution, but it embodies the essential features of

the distribution and makes it possible to look for evidence of these distribution maxima in the experimental echo-rate data. The discrete approximation is furthermore compensated for to some extent by the widths of the antenna patterns which tend to smooth the predicted diurnal variations.

The distributions of meteor activity around the earth's orbit is known to be non-uniform. However, if the form of the radiant distribution does not change appreciably around the earth's orbit, the computed pattern of the diurnal variations in meteor echo rate can simply be multiplied by an empirical factor which changes from month to month. The mean seasonal variations have been given by Hawkins (1956b), and his results are in agreement with the monthly variations in mean echo rate reported by Vogan and Campbell (1957) for the Greenwood–Ottawa path.

Comparison of Predicted and Observed Echo Rates

The exact computation of the diurnal variations in forward-scatter echo rate for a meteor shower has been described in Meeks and James (1957). Briefly, the procedure is to determine first the locus of points at a height h above the surface of the earth such that shower meteors will produce specular reflection between the transmitter and receiver. (The height h was set equal to 100 km for the calculations reported here.) Second, the relative effectiveness of points along the shower locus is determined independent of the antenna patterns of the transmitter and receiver. Third, the antenna illumination function is evaluated over the surface at a height h above the earth. This illumination function is $\sqrt{(G_t G_r)} \cos \mu$, where G_t and G_r are the respective gains of the transmitting and receiving antennae, and μ is the angle between the electric field of the illuminating wave and the electric field accepted by the receiving antenna. Finally the line integral of the product of the relative effectiveness and the antenna illumination function must be evaluated over the locus for specular reflection. This computation is then performed repeatedly at times spaced usually 1 hr apart throughout the day to obtain the pattern of diurnal variation in echo rate.

An example of the agreement which is found between the computed and observed echo-rate variations is shown in Fig. 1. This figure shows data obtained from an observation of the Geminid shower beginning at 1900 EST on 12 December 1957 and ending at 0300 EST on the following day. The observation was made on the 230-km forward-scatter path between Knoxville, Tennessee, and Atlanta, Georgia. A detailed description of this path is given in Table 1. The observation of the Geminid shower was made with the transmitting and receiving antennae directed toward a point 400 km south of the path midpoint and at a height of 100 km above the earth. The close agreement in Fig. 1 between the observed and predicted results for this short path tends to confirm the validity of the computational procedure.

The longest continuous series of meteoric forward-scatter observations has been reported in Vogan and Campbell (1957). These observations were made during a period of 15 months over the path from Greenwood, Nova Scotia, to Ottawa, Ontario, which is described in Table 1. The average hourly meteor rates from these data show wide fluctuations from day to day. However, the diurnal pattern

VOL. 16 1959

obtained by averaging over periods of 1 month is relatively smooth and shows small progressive changes from month to month. Fig. 2 reproduces the echo-rate data obtained by VOGAN and CAMPBELL for the even-numbered months. The

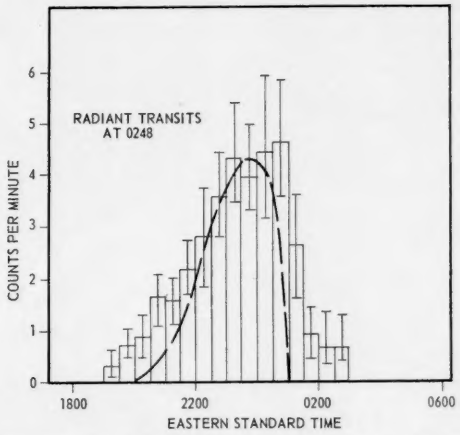


Fig. 1. Measured and predicted variations in meteor echo rate for the Geminid shower observed on the Knoxville–Atlanta path. The 95 per cent confidence limits are shown for the measured echo rates. The smooth curve shows the predicted variation in echo rate due to the shower.

scattered points for each hour represent the average echo rates for that hour on individual days. The monthly averages of these data are indicated by the solid lines which are drawn from one average hourly rate to the next.

Table 1. Characteristics of the meteoric forward-scatter paths

Path	Frequency (mc)	Transmitted power (kW)	Station separation (km)	Bearing of great circle path (E of N)	Latitude of path midpoint	Antenna types
Knoxville–Atlanta	41·94	0·5	230	13°	34·9°N	7-element Yagis
Greenwood–Ottawa	49·98	0·1	860	93°	45°N	5-element Yagis
Boston–Atlanta	49·44 73·82	5·0 3·0	1480	51·5°	38·2°N	5-element Yagis
Boston–Columbia	49·44 73·82	5·0 3·0	1250	44·1°	38·0°N	5-element Yagis

The histograms in Fig. 2 show the computed variations in the hourly echo rate based on the three-point-radiant approximation. The separate contributions from each radiant point are indicated in this figure.

The predicted diurnal pattern representing each month has been obtained by making the computation for the fifteenth day of the month (except in the case of June where the pattern is the average of 7 June and 21 June). The computed

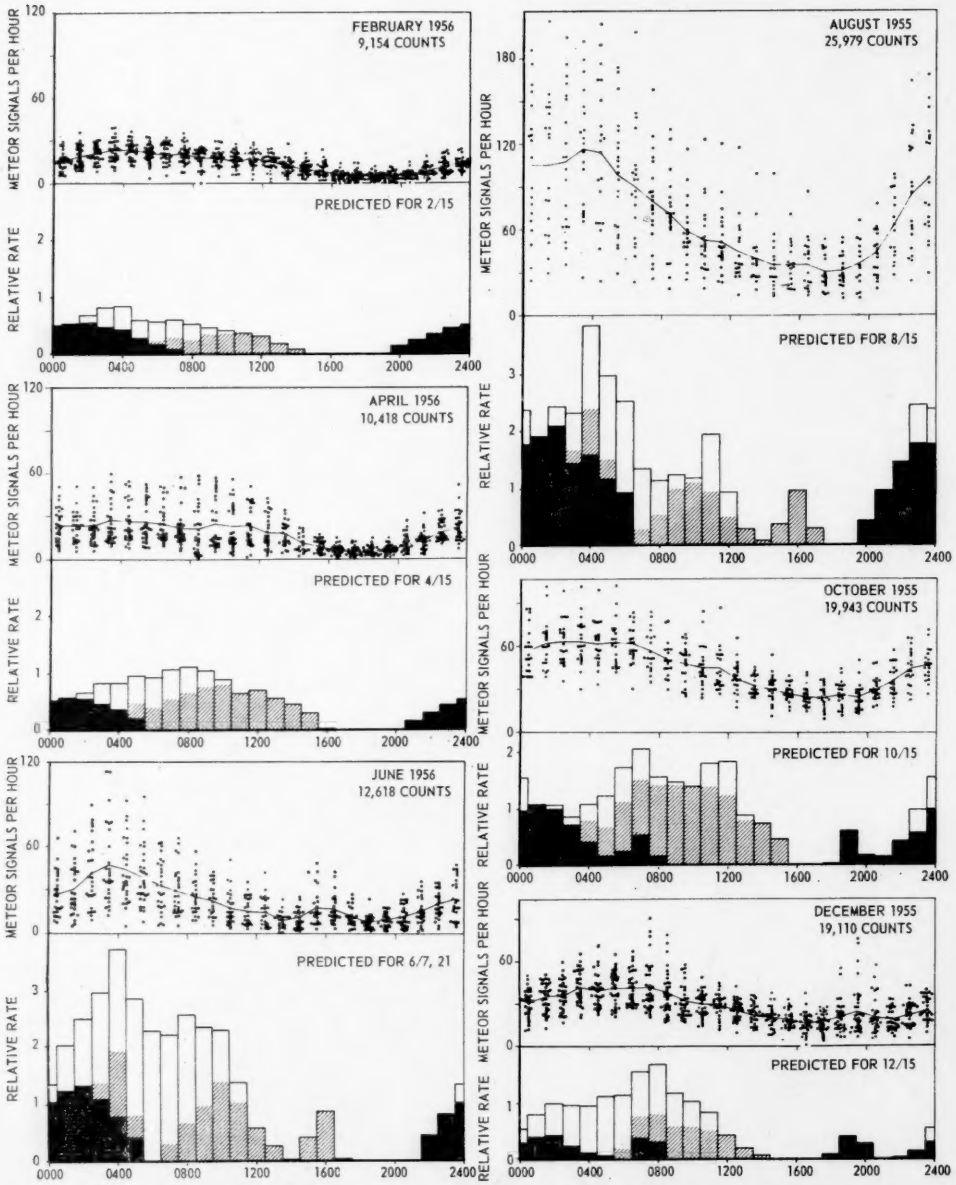


Fig. 2. Measured and predicted variations in meteor echo rate for the Greenwood–Ottawa path. Data reported by VOGAN and CAMPBELL (1957) are shown for the even-numbered months together with the predicted echo rate obtained from the three-radiant-point approximation. The separate contributions of the apex point (white), the sun point (grey) and the antihelion point (black) are indicated. All times are Eastern Standard.

VOL.
16
1959

diurnal distributions were multiplied by a factor given in HAWKINS (1956b) in order to take into account the non-uniformity of the density of meteor debris around the earth's orbit.

A comparison of the measured and computed diurnal rates shows surprisingly good agreement, if one considers the nature of the approximation. The prediction of zero echo rate in the afternoon hours results from the fact that all three of the point radiants in the approximate radiant distribution are below the horizon. It may be noted that the observed activity is a minimum at these times. In general the predictions for the 12 hr periods from noon to midnight are somewhat better than the predictions of the first half of the day. Of particular interest are the afternoon maxima predicted for June and December. These maxima, which are a result of the sun radiant point in June and the antihelion radiant point in December, appear clearly in the experimental data. The afternoon maxima predicted for August and October do not appear quite so clearly in the averaged data for these months. The comparatively sharp peaks at 0400 in June and August appear to be produced by the radiant concentrations near the sun. The 0800 peak in the December prediction as well as the 1900 peak are correlated with the antihelion concentration. The 0800 peak is a little less evident, however, in the December data than the 1900 peak. Generally the sun or antihelion concentration shows up best when it can produce peaks in the echo rate with the other concentrations below the horizon.

The sharp peaks in the diurnal rate curves which are predicted on the basis of the apex concentration do not appear, except possibly in October at 1130. However, in no case does the apex concentration contribute alone to the diurnal rate. When the apex is effective, the radiant distribution above the horizon is more complicated, and the three-point-radiant approximation appears to predict peaks and abrupt changes in the diurnal rate distribution which do not appear in the data. However, the representation of a continuous radiant distribution by three discrete radiant points may be expected to produce just this kind of discrepancy.

It is interesting to compare the predictions given here for the Greenwood–Ottawa path with predictions made in HINES (1956) for a similar path on the basis of a single diffuse distribution of radiants centred on the apex and symmetric about this point. Such a distribution is to be expected if the earth, during its motion around the sun, collides with meteors having a distribution of velocities which is isotropic in a heliocentric co-ordinate system. The earth's motion relative to the heliocentric system would then produce a diffuse concentration of meteor radiants toward the apex. When such a distribution is observed on an east–west, forward-scatter path, with identical antennae directed along the path, there is complete mirror symmetry between the geometrical situation at a time interval T before the apex transits the meridian of the path midpoint, and an interval T after this transit. The data in Fig. 2 do not in general show this symmetry, and consequently the apex distribution alone appears to be less representative of the actual distribution than the three-point-radiant approximation.

The average meteor echo rate has been measured during December 1957 and also during January and February 1958 over a path from Boston, Massachusetts, to Atlanta, Georgia. Echoes were recorded simultaneously at 49 Mc/s and 74 Mc/s.

Table 1 gives a description of this path and the equipment used in the measurements. The three-point-radiant approximation has been used to calculate the diurnal variations in the echo rate for this path. The average echo rate for each frequency is given in Fig. 3 together with the predicted rates for 12 December and 2 February. It is apparent in Fig. 3 that the predicted diurnal pattern agrees better with the data obtained at 74 Mc/s than at 49 Mc/s. However, many more counts were obtained at 74 Mc/s, where no ionospheric scatter appeared to obscure

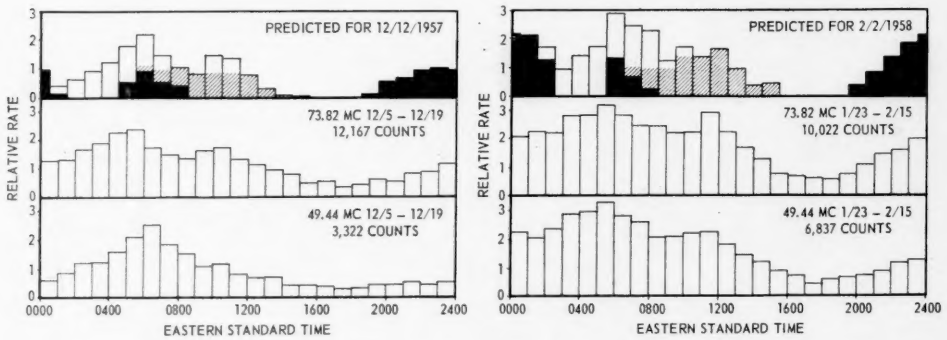


Fig. 3. Measured and predicted variations in meteor echo rate for the Boston–Atlanta path. The separate contributions predicted for the apex point (white), sun point (grey) and the antihelion point (black) are indicated. The left-hand portion of the figure shows the rates for December 1957 and the right-hand portion shows the rates for January and February, 1958. The observations at 74 Mc/s and 49 Mc/s were taken simultaneously.

VOL. 16 1959

the meteor echoes, and therefore the 74 Mc/s data can be expected to give a better estimate of the true echo-rate distribution. The peaks predicted for the diurnal distribution on 12 December show up in the average rates observed at 74 Mc/s. The 0600 peak in the computed rate is produced by the antihelion radiant point, and the 1000 peak is the combined result of the apex and sun radiant points. The predictions for 2 February include a peak at about 1200 due to the sun radiant point. This peak is observed at both frequencies. The contributions from the apex and antihelion points appear to coincide with the broad hump between 0500 and 0600. The predicted peak at midnight due to the antihelion radiant point does not stand out clearly.

Similar measurements were made over a path from Boston, Massachusetts, to Columbia, South Carolina. Table 1 gives a summary of the essential characteristics of this path. The average diurnal distributions of echo rates for two periods in the spring of 1958 are given in Fig. 4. Computed rates are also shown in this figure. The three sharp peaks predicted for 21 March fail to show up in the average diurnal distribution, and the observed distributions appear very much smoother than the predicted one. However, better agreement is found for the April data shown in Fig. 4. The broad maximum in the observed rate around 0800 corresponds to the combined activity of the apex and sun radiant points. The small maximum produced by the sun radiant point at 1400 also shows up in the average observed echo rates.

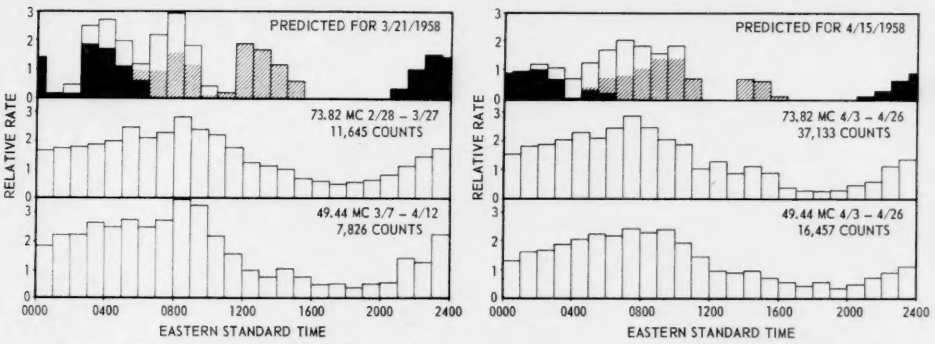


Fig. 4. Measured and predicted variations in meteor echo rate for the Boston–Columbia path. The separate contributions for the apex point (white), sun point (grey) and antihelion point (black) are indicated. The left-hand portion of the figure shows the average echo rates around March 1958 as observed on 74 Mc/s and 49 Mc/s. The right-hand portion shows average echo rates for April 1958.

Conclusions

The diurnal variations in the meteor echo rates observed by forward scattering appear generally to be consistent with the salient features of the distribution of sporadic meteor radiants in Hawkins (1956b). The concentrations of radiants near the sun and antihelion points should produce certain characteristic maxima in the diurnal echo-rate curves. Whenever these maxima are predicted for the afternoon hours, they are observed in the average diurnal-rate curves, as the graphs show in Fig. 2 (for the months of June, August, October and December) and Fig. 4 (for the period 3–26 April). During the hours before noon when the apex and one or both the other concentrations are above the horizon, the detailed shapes of the observed and predicted diurnal-rate curves do not always agree. This disagreement can perhaps be attributed to the choice of only three radiant points to represent Hawkins' continuous distribution of radiants. It is significant to note that the average diurnal-rate curves for periods of roughly 1 month show generally good agreement with the prominent features of a radiant distribution which was obtained by averaging about 2 years of radar data. Thus, in addition to confirming the principal features of Hawkins' average distribution of sporadic meteor radiants, the forward-scatter data indicates that this distribution leads to reasonably good predictions of the average diurnal variations in echo rate for periods as short as 1 month.

References

Forsyth P. A., Hines C. O. and Vogan E. L.	1955	*Canad. J. Phys.* **33,** 600.
Hawkins G. S.	1956a	*Mon. Not. R. Astr. Soc.* **116,** 92.
Hawkins G. S.	1956b	*Astr. J.* **61,** 386.
Hawkins G. S. and Prentice J. P. M.	1957	*Astr. J.* **62,** 234.
Hines C. O.	1955	*Canad. J. Phys.* **33,** 493.
Hines C. O.	1956	*J. Atmosph. Terr. Phys.* **9,** 229.
Hines C. O.	1958	*Canad. J. Phys.* **36,** 117.
Hines C. O. and Vogan E. L.	1957	*Canad. J. Phys.* **35,** 703.
Meeks M. L. and James J. C.	1957	*Proc. Inst. Radio Engrs.* **45,** 1724.
Vogan E. L. and Campbell L. L.	1957	*Canad. J. Phys.* **35,** 1176.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 236 to 245. Pergamon Press Ltd. Printed in Northern Ireland

The ratio of sodium to potassium in the upper atmosphere*

E. A. LYTLE and D. M. HUNTEN
University of Saskatchewan, Saskatoon, Saskatchewan, Canada

(*Received* 18 *December* 1958)

Abstract—Since sodium in the upper atmosphere is observable by the resonance scattering of the *D*-lines, it is natural to inquire whether potassium might also be found. Of the likely sources for these substances, sea-water has an atomic abundance ratio Na : K of 47 : 1, while the rest are in the range between 5 : 1 and 10 : 1. Consideration of the chemical and ionic equilibria suggests that the ratio of free, neutral atoms in the upper atmosphere should be about the same as in the source. The potassium resonance lines are at 7665 and 7699 Å; the former is the stronger but is obliterated by a strong line of the atmospheric *A*-band. An infra-red sensitive photomultiplier was used in a grating spectrometer along with a "condenser memory" permitting exposures to the twilight of 15 min or more. No potassium line has been detected; maximum possible intensities have been estimated from the tracings and for some occasions an upper limit of 30 : 1 has been achieved for the Na : K ratio. This suggests that the source is sea-water. Some measurements of the intensity ratio of the two *D*-lines are also reported.

1. INTRODUCTION

As soon as the presence of the sodium *D*-lines in the spectra of the twilight and night airglow had been established, the question of the source of the sodium arose. A good discussion of the matter was given by CHAPMAN (1939), and his conclusions still stand. The likely sources were: (a) sea-water; (b) volcanic dust; (c) meteorites; and (d) interplanetary matter. There is still very little evidence to bear on a choice between these, except that the volcanic dust seems unlikely because of the regularity from year to year of the intensity variation, especially in twilight. It would be useful to have a more detailed study of this from the southern hemisphere; it should show whether the variation correlates with the season or the time of year. The evidence available suggests the former, and thus points towards a terrestrial source; this would then probably be sea-water. A brief discussion by CHAMBERLAIN *et al.* (1958) shows that the observed amount of sodium in sea-level air is consistent with the amount in the upper atmosphere.

This paper points out that the ratio of sodium to potassium is much higher in sea-water than in any of the other possible sources. Though we have not been able to observe the twilight resonance of potassium, we have set a useful upper limit to its intensity, and thus to the abundance of potassium; the latter is about 3×10^8 atoms/cm^2 (column). The corresponding Na/K ratio is consistent only with the sea-water source unless our estimates of the factors which enter the calculation are wrong by a factor of 5 or so, and this is unlikely.

An attempt to observe the potassium resonance line was made by MEINEL a few years ago with an auroral spectrograph (SWINGS and MEINEL, 1952). Since no estimate was made of the intensity which would have been detected, it is impossible to compare the result with ours except to say that they are consistent.

* The research reported in this paper has been sponsored by the Geophysics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command, under Contract AF 19 (604)-1831.

VOL. 16 1959

During the investigation a number of measurements of the intensity ratio and absolute intensity of the two sodium *D*-lines were made. The results were not as good as had been hoped, and they are not strictly comparable with theory, but they are of some interest and also serve as a partial check on the estimates of the Na/K ratio. Both sodium and potassium measurements were made with a photoelectric scanning spectrometer aided by a condenser memory unit which allowed an "exposure" to the twilight of as long as 20 min and greatly increased the sensitivity of detection.

2. Expected Intensity of Potassium Lines in Twilight

The abundance ratios Na-atoms/K-atoms for various sources have been collected in Table 1. One would expect that the three values for various parts of

Table 1. Relative abundances (atoms/atom) of sodium to potassium in various sources

Source	Reference	Ratio Na/K
Cosmos	Kuiper (1954) p. 267	7·4
	Allen (1955) p. 28	10
Earth	Allen	3·25
Earth's crust	Allen	1·87
Earth's mantle	Allen	3·33
Sea-water	Sverdrup (1942) p. 168	47·2
Sun's photosphere	Unsöld (Kuiper (1953) p. 160)	12
	Strömgren (Kuiper (1953) p. 160)	4
	Claas (Kuiper (1953) p. 160)	20·9
Meteoritic silicate	Kuiper (1954) p. 267	6.6

the earth would be representative of volcanic dust, and those for the cosmos and the sun's photosphere of interplanetary matter; though of course the "cosmic" result has been derived primarily from the others. It may be seen that a ratio of about 7 is typical of these sources and also of meteorites, but that it is as high as 47 for sea-water; the latter is just beyond our detection, but the former should be observed easily. However, the ratio of free, neutral atoms in the upper atmosphere may not be the same as in the source, and it is necessary to consider the chemical and ionization equilibria.

Several papers have discussed the chemical equilibrium of sodium; the most recent is by Omholt (1957). The most important reactions are believed to be:

$$\mathrm{Na} + \mathrm{O_2} + \mathrm{M} \rightarrow \mathrm{NaO_2} + \mathrm{M}$$

$$\mathrm{NaO_2} + \mathrm{O} \rightarrow \mathrm{NaO} + \mathrm{O_2}$$

and

$$\mathrm{NaO} + \mathrm{O} \rightarrow \mathrm{Na} + \mathrm{O_2}.$$

M represents any third body. Because of the great chemical similarity of sodium and potassium, one would expect that just the same reactions would be important for the latter, and that the equilibrium would be similar. It is easy to write an

expression for the equilibrium between the three reactions, and for the ratio of free to combined metal, but the result depends on the reaction rates and also on the abundances of the other reacting bodies. But there is little doubt that most of the metal is free at high altitudes and combined at low; the height of the transition is uncertain but consistent with the observed maximum density of sodium at about 85 km. If the height were the same for potassium, the total abundances of free atoms would be in the same ratio as in the source. Since the reacting particles are the same, a difference could arise only from the reaction rates. The only guide we have is the semi-empirical rules of HIRSCHFELDER (1941) quoted by BATES and NICOLET (1950); they relate the reaction rate to the dissociation energy of the molecule which breaks up. Again, there is uncertainty in many of the dissociation energies for free molecules, but some are tabulated for the corresponding solids. As would be expected, they are nearly the same when sodium and potassium compounds are compared, and it would be surprising if the equilibrium constants we want should differ by more than a factor of 2. Because of the exponential variations in the other reacting bodies, this would make very little difference in the height of dissociation. Thus, a potassium layer should form at very nearly the same height as the sodium, with a peak within a few kilometers of 85.

The rate of ionization of sodium by solar ultra-violet has been calculated by BATES (1947) and corrected for more recent values of solar flux by HUNTEN (1954). The rate of recombination is uncertain, but again, should be very similar for both metals. The ionization cross-section of potassium has been measured by DITCHBURN *et al.* (1943) and may be combined with the solar flux given by TOUSEY (1953) to get the ionization rate. But the cross-section begins to rise towards shorter wavelength below 2700 Å, and the result would depend on the uncertain solar flux in this region if it were not for the absorption by O_2; this was represented by a cutoff at 2000 Å. We find the ionization rate to be $1{\cdot}6 \times 10^{-5}$/K atom per sec; this is almost equal to the corresponding 2×10^{-5}/Na atom per sec. Since the recombination rates should be closely equal also, it appears that the degree of ionization is essentially the same for sodium and potassium, or if anything, less for potassium.

The conclusion is that the ratio of free, neutral atoms in the upper atmosphere is equal to the atomic ratio in the source within a factor of 2. Since the abundance of sodium approaches 10^{10} atoms/cm^2 in winter, that of potassium should be about 10^9 atoms/cm^2 for three of the possible sources, and 2×10^8 for sea-water.

Next we must consider the expected brightness of potassium resonance radiation as a function of the abundance N (in atoms/cm^2). Reference to the Utrecht Atlas of the solar spectrum (MINNAERT *et al.*, 1940) shows that the stronger line at 7665 Å is obliterated by intense absorption line of the Fraunhofer A-band, which is due to atmospheric O_2. Fortunately, the other line, at 7699 Å, is free of this; but of course there is a solar absorption line of potassium. The atmospheric K will be excited by the residual intensity at the bottom of this; the Atlas shows about 30 per cent of the continuum, but this is certainly too high. An attempt was made to correct the profile for the resolution of the spectrograph by the method of BRACEWELL (1955), using the instrumental shape given at the beginning of the Atlas; the residual intensity came out at only a few per cent. It seems

VOL. 16 1959

unlikely that it would be smaller than that of the D-lines, about 5 per cent, and this result suggests that it is not much more; we adopt the value $r = (8 \pm 3)$ per cent. The flux at the earth in the nearby continuum can be found from the tables of MINNAERT (1953) as $f'_\nu = 930$ quanta/cm^2 sec sec^{-1}.

For the abundances we expect to find, there will be negligible self-absorption; hence the brightness is given by

$$4\pi B = r f'_\nu \left(\frac{\pi e^2}{mc}\right) N f \,.\, 10^{-6} \text{ Rayleighs (R)}.$$

The unit is defined by HUNTEN *et al.* (1956); the quantities inside the brackets have their standard meanings; and f represents the oscillator strength which is 0·33 (ALLEN, 1955). Therefore:

$$N = 1{\cdot}5 \times 10^6 \,.\, 4\pi B \text{ atoms/cm}^2;$$

we hope that this should be correct within 50 per cent. From the estimates of abundance given already, we find that the expected intensity in winter is 130 R for sea-water and 650 R for the other sources.

3. Observational Technique

The detecting instrument used was the auroral scanning spectrometer (HUNTEN 1953) which has been very successful in observing sodium emission; in this service, it can detect about 200 R in a 1 min scan. The photomultiplier still has a slight sensitivity at 7699 Å, but certainly not enough to be useful for twilight work. A tube with a Cs–O–Ag cathode is more suitable, and the one used was the experimental DuMont K1292. It is well known that this surface has a high dark current; it was reduced by refrigeration with dry ice. It may be that a considerable further improvement might have been possible by the use of liquid nitrogen, but this was not tried. At first it was thought that it might be sufficient to cool only the cathode end of the bulb, leaving the base at room temperature, but this did not work; cooling the whole tube was much better.

The infra-red tube is much less sensitive than the familiar types with the Sb-Cs cathode; the minimum detectable intensity at 7699 Å is about 4000 R. By observing at the horizon, a gain of about 4 may be obtained, but this is still not enough. If the scanning time could be increased from 1 to 16 min, with a corresponding increase in the filtering time-constant, a further gain of 4 would be possible, to about 1000 R at the horizon or 250 R referred to the zenith. This would be just adequate to distinguish between the possibilities. It is not practical to make a single 16 min scan on twilight, because the intensity would vary too much during the scan, but the same effect may be had by averaging (or adding together) sixteen scans of 1 min each. This was done automatically by a thirty channel condenser "memory" unit (HUNTEN, 1959). Each spectrum is sliced up into thirty parts, and the parts are stored by low-leakage condensers. As a new spectrum comes in it is added to the sum already present. Typical results may be seen in Figs. 2 and 3.

To get the longest possible exposure to the highest possible intensity, a chosen shadow-level of the twilight was followed around the horizon from an azimuth at

right angles to the sun to one above the sun. Directions away from the sun were not included because they give a higher background intensity from the solar continuum and make the line more difficult to detect. When a shadow height has been assumed, a programme may be worked out giving the required azimuth (with respect to the sun) as a function of solar depression. This is good for any twilight; the corresponding times and solar azimuths may be found for a given night with the aid of a nomogram (HUNTEN, 1958).

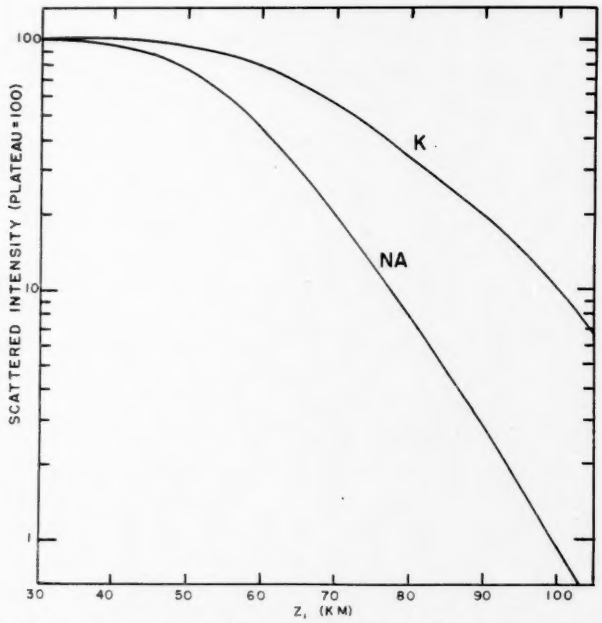


Fig. 1. Twilight intensity curves for Na and K, calculated for layers peaked at 85 km. z_1 represents the geometrical shadow height.

To choose a suitable shadow height, it is necessary to know how the line intensity should vary during twilight. The calculations were made in the same way described by HUNTEN (1954) assuming a symmetrical distribution of K with a peak at 85 km as for Na. The result is shown in Fig. 1 along with a sodium curve for comparison. The potassium curve does not begin to drop until considerably later; this is because the shadowing by ozone is negligible, and to a smaller extent because the Rayleigh extinction is lower. Usually it is important to have the highest possible ratio of emission line to continuum, and this will be obtained just after the line intensity starts to fall: about 50 km for Na and 60 km for K; the corresponding depression angles (at the point observed) are 7°09′ and 7°49′. It might seem that the sensitivity could be increased by using wider slits to pass more light, but this is not so if the noise is controlled by the photo-current from the continuum. Doubling the width gives twice the line signal but four times the white signal and therefore twice the noise; the signal/noise ratio for the line is unchanged. However, a gain does result if the noise is controlled by the dark

VOL. 16 1959

current. On the other hand, the wider the slits, the more uncertain is the identity of any line that may be observed. For K on most occasions a resolution of 20 Å was used, but sometimes 10 Å was tried. For Na it was maintained at 3 Å because the two lines were to be resolved for ratio measurements. The memory was also used here to improve the accuracy, even though the lines could be observed without it.

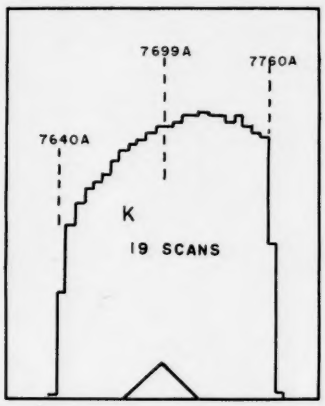


Fig. 2. A spectrum of twilight in the region of the potassium resonance line. The steps are introduced by the channels of the condenser memory. Resolution is such that a line would give the triangle at the bottom.

To estimate the abundance it is necessary to know the plateau intensity, but our observations were made just off the plateau. They were corrected by factors taken from Fig. 1; this procedure introduces further errors, but it could not be helped since the highest possible signal/background ratio was needed for both K and Na.

4. Potassium: Results and Discussion

Though the greatest intensity of the potassium line was expected in winter, observations were carried out for most of a year, from the summer of 1957 to the spring of 1958. On no occasion was the line detected, but limits of detection were estimated from the tracings (such as Fig. 2) and the results for winter are given in Table 2. To compare with expectations, they must be referred to the zenith and corrected to the plateau; the latter has just been discussed. A layer at 85 km and observed at a zenith angle of 80° is about four times brighter than in the zenith; this includes a slight reduction for extinction. The spectrum shown in Fig. 2 is typical of the better ones; the marked depression at the short-wavelength end is caused by the unresolved lines of the Fraunhofer *A*-band, one of which blots out the other potassium line.

It was not feasible to measure the sodium brightness on the same occasions, since no other instrument was available for the purpose. Instead, the average results for several previous years were used (Chamberlain *et al.*, 1958) and checked occasionally by the sodium runs reported in the next section. There seems to be no evidence that this year was different from previous ones. These

average results, combined with the estimates for potassium, give the lower limits to the Na/K ratio in the last column of Table 2.

Some of these approach the 47 expected for the sea-water source, and all are greater than the values of around 7 for the other sources. As is pointed out in Section 2, there are uncertainties in the chemical equilibrium, in the degree of ionization and in the residual intensity of the Fraunhofer line. These errors

Table 2. Summary of the potassium observations made when the sodium abundance was large

Date 1957–1958	Shadow height (z_1) (km)	$4\pi B$ U.L. (R)	$4\pi B$, corr to zenith and plateau, U.L. (R)	N(K) U.L. (10^8atoms/cm^2)	N(Na) 10^9 atoms/cm^2	Na/K L.L.
23 Nov. p.m.	60	492	164	2·5	4·3	17
2 Dec. p.m.	60	322	107	1·6	4·6	29
3 Dec. p.m.	60	713	237	3·5	4·6	13
4 Dec. a.m.	60	924	308	4·6	4·6	10
10 Jan. a.m.	60	1130	386	5·6	5·7	10
15 Jan. a.m.	60	920	307	4·6	6·0	13
20 Jan. p.m.	60	699	233	3·5	6·2	18
25 Jan. p.m.	50	912	256	3·8	6·3	17
2 Feb. p.m.	40	892	232	3·5	6·3	18
12 Feb. a.m.	60	364	121	1·8	6·3	35
14 Feb. p.m.	60	1030	343	5·1	6·4	13
21 Feb. p.m.	60	380	127	1·9	6·4	34

U.L. means upper limit and L.L. means lower limit.

VOL. 16 1959

would affect all the results in the same way, and may be called "systematic"; the error in the sodium abundances is random and will probably average out in a number of observations as large as this. It would seem unlikely that the errors are all in the same direction and large enough to make the results consistent with a Na/K ratio as low as 7 or 10, though this possibility cannot be completely ruled out. It may be concluded that probably the ratio is about 50 and the source of upper-atmospheric sodium is evaporated sea-spray. This is supported by the other evidence mentioned in Section 1.

5. Sodium: Results and Discussion

The sodium measurements were undertaken for two purposes: to compare with the potassium, and to provide simultaneous measurements of the intensity and the ratio of the two lines. According to the discussion of Chamberlain *et al.* (1958), these quantities have been measured separately before, but not on the same occasion; the latter would give a more useful check on twilight theory. Montalbetti (private communication) has many measurements of intensity through twilight in the two *D*-lines separately, and these should give a more detailed check; but the present results are worth a brief consideration.

In the summer, the *D*-lines are faint and it was necessary to use the memory unit to get a sufficient signal/noise ratio. For the best measurement of the intensity

and the best comparison with theory, the measurements should be made on the plateau, but it was found that the background continuum was too high there. The intensities must be measured from the bottom of the Fraunhofer line which appears in early twilight under the same measuring conditions, as illustrated in Fig. 3; if the continuum is very strong, the amount of this correction gives the major uncertainty in the D_2/D_1 ratio. Thus the observations were programmed for a height of 60 km, just off the plateau, and corrected from Fig. 2. The zenith angle was 75°. In winter the intensity is much higher, and there is no difficulty in making measurements; on a few occasions two runs were made at different

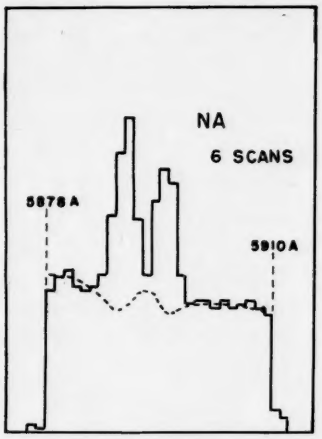


Fig. 3. A spectrum of the D-lines suitable for measuring the intensity ratio; measurements must be made from the continuum shown by a dotted line. Good accuracy is attained by measuring the areas of the lines, not their peaks.

heights to check the correction to the plateau. Fig. 3 is from one of these; an excellent spectrum was obtained in only 6 min. We believe that the measurements of the ratio are accurate to a few per cent.

A reasonable way to compare the results with theory is to calculate the abundance from the intensity by means of the curve given by CHAMBERLAIN *et al.* (1958) and to plot the observed ratio against this abundance. This may then be compared with the theoretical curve of ratio vs. abundance. The result was unsatisfactory with the points clustering well below the line; our interpretation of this is that the intensities are often lower than they should be and the abundances lower still. This is not unreasonable since the observations were made around the horizon at a zenith angle of 75° and must often have suffered considerable extinction.

A more satisfactory plot was obtained by using for the abundance the greater of: (a) the observed; or (b) the average given by CHAMBERLAIN *et al.* from previous work. Since it is of little quantitative value, it is not reproduced, but we may draw one or two qualitative conclusions from it. First, it gives evidence that on many occasions the measured intensities at the horizon are indeed too low; and second, the fluctuations in the ratio are still larger than would be expected from the uncertainty in the abundance. For example, near the abundance

$6{\cdot}5 \times 10^9$ atoms/cm^2 the ratios vary from 1·54 to 1·12; only the first mentioned is above the predicted 1·43, and only two points in all lie above the curve. Just this fluctuation has been predicted by DONAHUE (1958) who has found that several effects can seriously perturb the ratio when the sodium layer is in partial shadow as it was for most of these observations. Moreover, the predictions we have been using were made for the time just before the shadow reaches the layer and are not properly comparable.

Thus our main conclusion must be that more observations of intensity ratio are necessary, and that they should be made in such a way as to satisfy the conditions adopted for theoretical calculations. This should be possible in winter but probably difficult in summer. The fact that the ratios have a seasonal variation of the type predicted suggests that the sodium abundance followed its usual course during the year under consideration, as has already been assumed in discussing the search for potassium. On some occasions, the potassium emission may have been weakened in the same way as the sodium, but probably about half the observations are not seriously affected. Thus our conclusions about the Na/K ratio remain unchanged.

6. CONCLUSIONS

It has been shown that it should be possible to distinguish between two groups of possible sources for upper-atmospheric sodium by measuring the ratio of sodium to potassium emissions in twilight. Sea-water has a Na/K ratio of 47, and the others between 5 and 10. A discussion of the chemical equilibrium and state of ionization has shown that potassium should exist in the free, neutral form to approximately the same extent as sodium, and should thus be found in a layer near 85 km. The failure to observe any sign of the potassium resonance line at 7699 Å has shown with little doubt that the source must be sea-water.

Some measurements of the intensity and intensity ratio of the two sodium *D*-lines have also been made. The observed intensities are apparently not reliable, but the ratios suggest that the abundances for this year were normal. Large fluctuations from the average were found for the ratio.

A technique for accumulating the greatest possible exposure to a twilight spectrum has been described. The use of a condenser memory unit has allowed this to be used with a photoelectric scanning spectrometer. Although it does improve the sensitivity of detection, it leads to difficulties of interpretation and should not be used unless there is no alternative.

Acknowledgements—We are grateful to Dr. R. MONTALBETTI for discussions of our own and his results on the *D*-lines; and to Dr. T. M. DONAHUE for discussion of their interpretation.

REFERENCES

ALLEN C. W.	1955	*Astrophysical Quantities.* Athlone Press, London.
BATES D. R.	1947	*Terr. Magn. Atmos. Elect.* **52,** 71.
BATES D. R. and NICOLET M.	1950	*J. Geophys. Research* **55,** 301.
BRACEWELL R. N.	1955	*J. Opt. Soc. Amer.* **45,** 873.
CHAMBERLAIN J. W., HUNTEN D. M. and MACK J. E.	1958	*J. Atmosph. Terr. Phys.* **12,** 153.

VOL. 16 1959

Chapman S.	1939	*Astrophys. J.* **90,** 309.
Ditchburn R. W., Tunstead J. and Yates J. A.	1943	*Proc. Roy. Soc.* A **181,** 386.
Donahue T. M.	1958	*Phys. Rev. Letters* **1,** 287.
Hirschfelder J. O.	1941	*J. Chem. Phys.* **9,** 645.
Hunten D. M.	1953	*Canad. J. Phys.* **31,** 681.
Hunten D. M.	1954	*J. Atmosph. Terr. Phys.* **5,** 44.
Hunten D. M.	1958	Scientific Report No. BR-8, Contract AF 19 (604)-1831, University of Saskatchewan, Saskatoon.
Hunten D. M.	1959	*Canad. J. Phys.* In preparation.
Hunten D. M., Roach F. E. and Chamberlain J. W.	1956	*J. Atmosph. Terr. Phys.* **8,** 345.
Hunten D. M. and Shepherd G. G.	1954	*J. Atmosph. Terr. Phys.* **5,** 57.
Kuiper G. P.	1953	*The Sun.* University of Chicago Press.
Kuiper G. P.	1954	*The Earth as a Planet.* University of Chicago Press.
Minnaert M.	1953	In Kuiper, 1953.
Minnaert M., Mulders G. F. W. and Houtgast J.	1940	*Photometric Atlas of the Solar Spectrum.* Utrecht Observatory.
Omholt A.	1957	*J. Geophys. Res.* **62,** 207.
Sverdrup H. U.	1942	*The Oceans.* Prentice-Hall, New York.
Swings P. and Meinel A. B.	1952	*The Atmospheres of the Earth and Planets* (Edited by Kuiper G. P.) (2nd. Ed.). University of Chicago Press.
Tousey R.	1953	*J. Opt. Soc. Amer.* **43,** 245.

L.
6
59

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 246 to 251. Pergamon Press Ltd. Printed in Northern Ireland

The infra-red spectrum of the night airglow from 1·4 μ to 4·0 μ*

J. F. Noxon, A. W. Harrison† and A. Vallance Jones
Physics Department, University of Saskatchewan,
Saskatoon, Saskatchewan, Canada

(*Received* 18 *December* 1958)

Abstract—The emission from the night airglow in the 1·4 μ to 4·0 μ region has been studied with a new infra-red spectrometer. The 9,7 OH vibration–rotation band at 2·16 μ has been observed for the first time and its relative brightness has been found to be somewhat lower than has been predicted theoretically. Beyond 2·5 μ thermal radiation from the lower atmosphere becomes too bright to allow the detection of further OH bands. The absolute brightness of the spectrum has been measured.

1. Introduction

In recent years the night airglow spectrum has been examined out to 2·0 μ with the help of an infra-red spectrometer equipped with a PbS photoconductive detector (Gush and Vallance Jones, 1955; Harrison and Vallance Jones, 1957). This work has shown that in the 1·0 to 2·0 μ region the night airglow emission consists principally of further bands of the OH vibration–rotation system discovered by Meinel (1950a) in the photographic infra-red. On the basis of the known energy levels of the ground state of OH, it was to be expected that more bands might be found beyond 2·0 μ. In particular there should be the 9,7 band at 2·16 μ and the entire $\Delta v = 1$ sequence extending from the 0,1 band at 2·80 μ to the 9,8 band at 4·46 μ. The possibility of observing these bands is determined by the sensitivity of the spectrometer, the transmittance of the atmosphere and the magnitude of the background thermal emission from the lower atmosphere. While a spectrometer of adequate sensitivity could be realized and the transmittance of the atmosphere is adequate at least in the 2·0–2·5 μ and 3·0–3·9 μ windows, there seemed to be a strong possibility that the thermal emission of the lower atmosphere might overwhelm the OH bands at longer wavelengths. The investigation was, however, pursued with the object of finding which OH bands could be observed and of studying the thermal emission itself.

2. Instrumentation

The basic design of the new spectrometer is similar to that described by Gush and Vallance Jones (1955). The incoming radiation is modulated at 135 c/s by the chopper which is cooled by dry ice to minimize radiation from its blades. The radiation is dispersed by a 300 lines/mm, 4 × 5 in. grating blazed at 1·9 μ and is finally focussed on a PbS cell by an ellipsoidal mirror. A cam and lever arrangement allows the grating to scan over any desired wavelength range with a

* The research reported in this paper has been sponsored by the Geophysics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command, under Contract AF 19 (604)-1831.

† Present address: Canadian Armament Research and Development Establishment, Valcartier, P. Q., Canada.

VOL. 16 1959

period of 3 min. The detector, an Infrared Industries PbS cell, is cooled by liquid air to improve its long wavelength response. The effective field of view of the spectrometer is approximately 10° square.

3. Results

Fig. 1 shows the night sky emission spectrum from 1·4 to 4·0 μ obtained at a zenith angle of 80° with a spectral slit width of 300 Å (0·03 μ). This spectrum was recorded during a very clear night on which the thermal emission from the lower atmosphere should have been relatively low since the ground air temperature

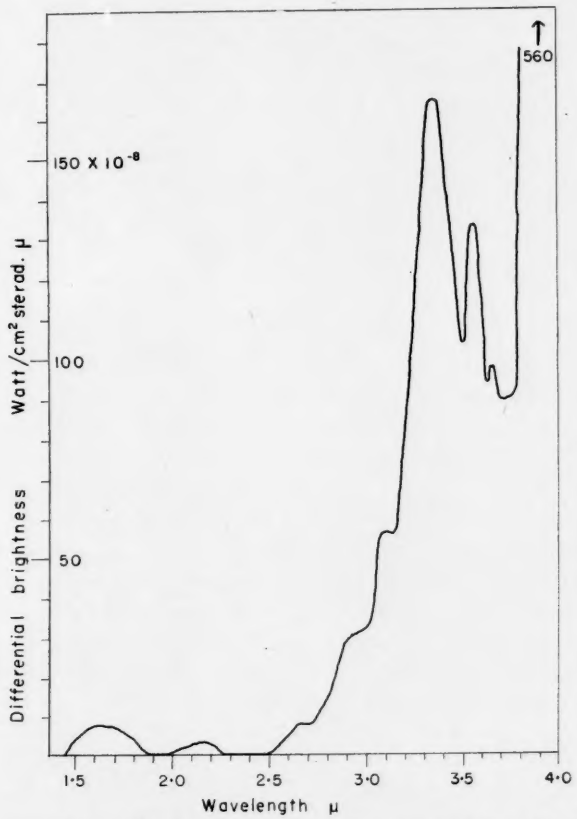


Fig. 1. Spectrum of the night sky from 1·4 μ to 4·0 μ; spectral slit width 300 Å; zenith angle of observation 80°; ground air temperature 252 $\pm$ 3°K.

was 252° $\pm$ 3°K. The absolute values of differential brightness were obtained by calibrating the spectrometer with a blackbody source.

The two peaks at 1·6 μ and 2·2 μ are due to OH. The former is that first observed by Gush and Vallance Jones (1955) and is a blend of the 3,1; 4,2 and 5,3 bands; the latter, observed here for the first time, is to be attributed principally to the 9,7 band at 2·16 μ since the adjacent 8,6 band at 2·00 μ is strongly

attenuated by atmospheric absorption, as are the shorter wavelength 7,5 and 6,4 bands. The emission between 2·5 μ and 4·0 μ will be shown in a following section to be predominantly thermal radiation from the lower atmosphere and to be considerably in excess of the predicted OH intensity. Consequently, observations of the $\Delta v = 1$ sequence of OH from low altitudes under normal atmospheric conditions are precluded.

4. Discussion

4.1. OH *emission bands*

The integrated OH brightness between 1·0 and 2·0 μ comes to $2{\cdot}4 \times 10^{-1}$ erg/cm^2 sterad-sec. When this value is referred to the zenith by dividing by a Van Rhijn factor of 4·0, a value of $6{\cdot}0 \times 10^{-2}$ erg/cm^2 sterad-sec is obtained. This value agrees well with the value of $5{\cdot}5 \times 10^{-2}$ erg/cm^2 sterad-sec obtained by Harrison and Vallance Jones (1957) who used a different instrument, calibrated independently.

Heaps and Herzberg (1952) have calculated approximate values of the transition probabilities among the lower levels of the ground state of OH. The main uncertainty in this calculation arises from a lack of knowledge of how the electric dipole moment of the molecule varies with internuclear distance. Heaps and Herzberg finally obtained a set of relative intensities from the OH emission bands by making the simplifying assumption that OH is excited exclusively into the ninth vibrational level. The relative intensities so calculated agree fairly well with the results of Meinel for the photographic infra-red region and with those of Harrison and Vallance Jones for the 1·0 to 1·5 μ region. The predicted emission intensities were compared with results of Fig. 1 by correcting the calculated intensities for the strong atmospheric absorptions centred at 1·4 μ and 1·9 μ with the help of the data of Taylor and Yates (1957); the integrated emission for the 1·6 μ peak should then be about 1·2 times stronger than the 2·16 μ peak. The observed ratio, obtained from Fig. 1 is about 4·0. Since 80 per cent of the 2·16 μ emission is due to the 9,7 band it would appear that the brightness of this band is significantly lower than predicted, relative to other bands of the same sequence with lower v.

This divergence between theory and observation could arise either as a result of the approximations involved in the calculations of the transition probabilities or from the assumption that OH molecules are formed exclusively in the $v = 9$ vibrational level.

Further evidence that the bands originating from the $v = 9$ vibrational level are weaker than predicted is provided by observations of the relative intensity of the 9,5 and 4,1 bands made by Fedorova, (1957) and Vallance Jones (1955). According to the predictions of Heaps and Herzberg the 9,5 band should be three times brighter than the neighbouring 4,1 band. As was pointed out by Fedorova however, the observations show that the 9,5 band is actually weaker than the 4,0 band. Heaps and Herzberg found also that the 9,4 and the 9,3 bands were somewhat weaker than predicted.

There is therefore, some evidence to support the contention that the initial excitation is not confined to the ninth vibrational level although at present the

VOL. 16 1959

results might alternatively be interpreted as indicating that the calculated transition probabilities are too large in the region of high v.

In Section 4·2 it will be shown that the emission observed beyond 2·5 μ can be accounted for mainly by thermal radiation from the lower atmosphere. In addition it can be shown that the OH bands of the $\Delta v = 1$ sequence, lying between 2·8 and 4·5 μ, do have a predicted intensity which is small compared with the night sky emission actually observed in this spectral region. This is most conveniently done by calculating a synthetic spectrum for the OH bands for a spectral slit width equal

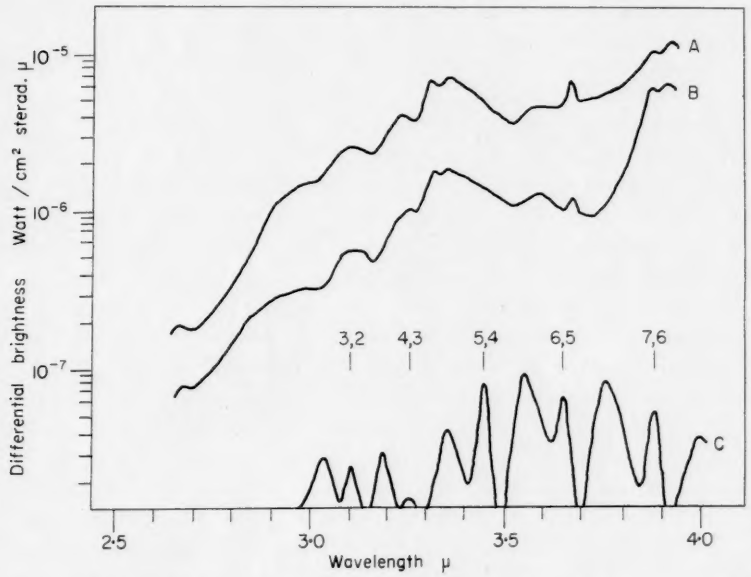


Fig. 2 Night sky spectra 2·6 μ–4·0 μ for a zenith angle of 80° and a spectral slit width of 300 Å. *A*: observed spectrum, ground air temperature 262°K. *B*: observed spectrum, ground air temperature 252°K. *C*: approximate predicted synthetic spectrum for OH emission.

to that actually employed in obtaining the spectrum shown in Fig. 1. An approximate synthetic spectrum for a spectral slit width of 300 Å was calculated, for a rotational temperature of 260°K, by combining the predicted relative intensities given by Heaps and Herzberg with the absolute intensity of 175 kilorayleighs (KR) given by Harrison and Vallance Jones for the zenith intensity of the 4,2 band. The resulting spectrum was roughly corrected for atmospheric absorption using the information given by Taylor and Yates (1957) and the intensities finally multiplied by 4 to give the predicted OH spectrum as it would appear at a zenith angle of 80°. The synthetic spectrum so obtained is plotted logarithmically as curve *C* in Fig. 2; the central *Q* branch is indicated for each band and the intervening peaks are the superposition of the *P*- and *R*-branches of neighbouring bands. Curve *B* is a logarithmic plot of the observed night sky spectrum given in Fig. 1 and curve *A* is a similar spectrum obtained on a warmer night. It is clear

that the observed night sky spectrum shows none of the structure to be expected if OH were a major contributor; furthermore the emission actually observed exceeds the predicted OH intensity by fully an order of magnitude. Hence the greater part of the emission beyond 2·5 μ must originate from a source other than OH.

4.2. *Thermal emission from the lower atmosphere*

The supposition that the emission observed between 2·5 and 4·0 μ is predominantly of thermal origin may be tested in several ways. If the lower atmosphere were isothermal then it should radiate as a black body of which the emissivity would equal the atmospheric absorbance at each wavelength, according to Kirchhoff's law. Consequently the thermal emission of an isothermal atmosphere should be given by the product of the corresponding perfect black-body emission and the fractional atmospheric absorption. An effective atmospheric temperature may be obtained by noting that the atmospheric absorption approaches 100 per cent in the vicinity of 2·7 μ; the observed night sky emission at this wavelength, as given in Fig. 1, leads to a black-body temperature of 258°K; this is fairly close to the ground air temperature which was 252°K when the spectrum was obtained. Using values for the atmospheric absorption given by Taylor and Yates a synthetic atmospheric emission spectrum was calculated for a spectral resolution of 300 Å, a value equal to the spectral slit width of the spectrometer.

The results are shown in Fig. 3 where curve A is the emission of 258°K black-body, curve B is the synthetic spectrum, and curve C is the observed night sky spectrum. Comparison between curves B and C shows a general correspondence; quantitative agreement is not be to expected because the atmosphere is not isothermal and because the actual absorbance curve for the atmosphere along the line of sight of the spectrometer almost certainly should be different from that given by Taylor and Yates who used a 3 miles absorption path along the ground. The atmospheric constituents responsible for absorption in the 2·5–4·0 μ region are also indicated on Fig. 3 following Shaw *et al.* (1951). The increased atmospheric absorption at 3·67 μ and 3·90 μ due to HDO and N_2O shows up as an increase in the thermal emission at these wavelengths.

A further check is provided by comparing the two night sky spectra shown as curves A and B in Fig. 2; the spectra were obtained on two different nights when the ground air temperatures were 262°K and 252°K respectively. If the emission is of thermal origin, then a rise in the overall atmospheric temperature should result in increased emission at all wavelengths with no pronounced change in the form of the spectrum. A comparison of curves A and B in Fig. 2 bears out this prediction, although the enhancement appears to be somewhat greater in the range 2·9 μ to 3·8 μ. Reference to Fig. 3 shows that in this wavelength region the atmospheric absorption is due primarily to H_2O and HDO; hence the additional enhancement probably resulted from an increase in the absolute humidity on the warmer night. Using the measured intensity at 2·7 μ the effective black-body temperatures associated with the two spectra are 265°K and 258°K respectively; the difference appears quite consistent with the change in the ground air temperature.

VOL. 16 1959

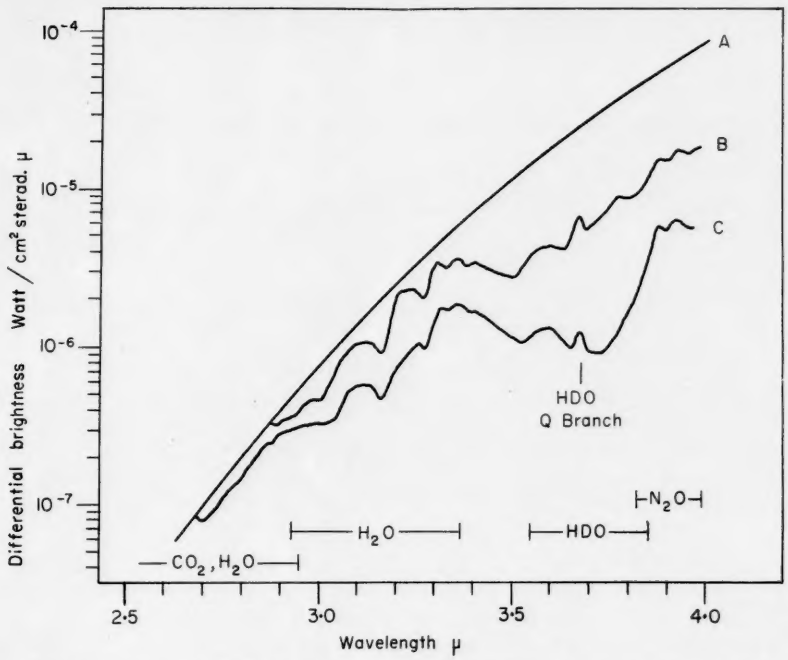


Fig. 3. *A*: 258°K black-body spectrum; *B*: "synthetic" atmospheric thermal emission spectrum. *C*: observed atmospheric thermal emission spectrum. Spectrum *C* is a logarithmic plot of the spectrum of Fig. 1. The atmospheric constituents principally responsible for absorption are indicated.

5. Conclusions

In addition to revealing the 9,7 OH bands at 2·16 μ in the night airglow the investigation has shown that thermal emissions from the lower atmosphere dominate the airglow spectrum at wavelengths greater than 2·5 μ. It did not prove possible to observe any bands of the $\Delta v = 1$ sequence of OH in the presence of the thermal radiation.

Acknowledgement—The authors are grateful to Mr. F. Rittmann for his fine workmanship in the construction of the new spectrometer.

References

Fedorova N. I.	1957	*Astron. Zh.* **34,** 247.
Gush H. P. and Vallance Jones A.	1955	*J. Atmosph. Terr. Phys.* **7,** 285.
Harrison A. W. and Vallance Jones A.	1957	*J. Atmosph. Terr. Phys.* **11,** 192.
Heaps H. S. and Herzberg G.	1952	*Z. Phys.* **133,** 48.
Meinel A. B.	1950a	*Astrophys. J.* **111,** 555.
Meinel A. B.	1950b	*Astrophys. J.* **112,** 120.
Shaw J. R., Chapman R. M., Howard J. N. and Oxholm M. L.	1951	*Astrophys. J.* **113,** 268.
Taylor J. H. and Yates H. W.	1957	*J. Opt. Soc. Amer.* **47,** 223.
Jones A. Vallance	1955	*Nature, Lond.* **175,** 950.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16. pp. 252 to 258. Pergamon Press Ltd. Printed in Northern Ireland

Auroral and nightglow observations at Ås, Norway

G. KVIFTE
Department of Physics, Agricultural College of Norway

(*Received* 20 *December* 1958)

Abstract—A new grating spectrograph has been used in the study of the upper atmosphere. A summary is given of the results of auroral observations. In a more detailed account of nightglow investigations are given precise wavelengths of the band lines of the 5,0, 6,1, 8,2 and 9,3 bands of the Meinel OH system, and absolute intensities of the bands. From the intensity distribution of the lines within the 6,1 and 9,3 bands is derived a rotational temperature of the emitting layer of 215°K. A sharp H_α line and the red forbidden lines of *NII* are found to be present in the nightglow.

INTRODUCTION

As part of the current investigations of the upper atmosphere in Norway, a grating spectrograph has been constructed at the Department of Physics, Agricultural College of Norway.

The design of the spectrograph is conventional. The collimated light from a Hilger symmetrical slit, type-F 1386, is reflected by a Bausch and Lomb plane grating, 76 × 65 mm², 2160 lines/mm, blazed for maximum efficiency at 6000 Å in the first order. The spectrum is photographed by means of a "Wray"* objective F/0·71, $f = 64$ mm. The overall linear dispersion of the spectrograms is about 35 Å/mm, and the spectral range covered, at a fixed position of the grating, is about 600 Å. When in use, a portion of the sky was projected on the vertical slit by a lens. The spectrograph was mounted on a hinged plate of a roller table and placed in a tower on a five-storied building at the college. Observation was possible through open windows for all azimuths and for zenith angles between about 30° to 90°. The geographical co-ordinates of the place of observation is 10°47′ E Gr, 59°40′ N (about 30 km south of Oslo) and 110 m above sea level.

The spectrograph was finished in the autumn of 1957, and during the winter a number of spectrograms, covering the range from 5150 Å to 6650 Å was obtained from aurorae and nightglow. Although the great light-power of the camera lens was not fully utilized, exposure times from 4 to 8 hr were sufficient to obtain good nightglow spectrograms. Fig. 1, a negative print of a nightglow spectrogram from the red region, should show the quality of the spectrograph.

POLAR AURORAE

The auroral spectrograms are mostly from relatively weak displays, and the exposure times have therefore always been sufficiently long for the OH Meinel bands to appear. They do not, however, seem to be enhanced by the auroral activity.

The O_2^+ first negative and the N_2 first positive bands were observed with

* From Wray Optical Works Ltd., England.

VOL. 16 1959

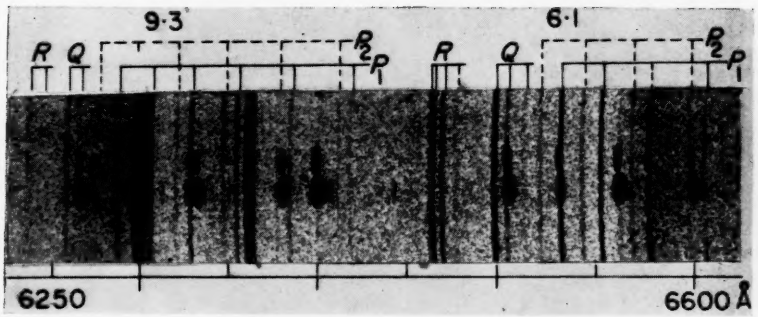


Fig. 1. Nightglow spectrogram showing the 6,1 and 9,3 OH bands.

VOL.
16
1959

variable intensities. The 4,0 and 5,1 bands of the N_2^+ Meinel system appeared weakly on one spectrogram.

All O*I* lines of the normal triplet and quintet series, lying in the region from 5300 Å to 6500 Å, are definitely present, including the multiplets at 5958 Å and 6046 Å, reported masked or absent by PETRIE and SMALL (1952). The mean intensities of the multiplets mentioned, measured from two auroral spectrograms, lie between a half and a third of the intensity of the stronger component of the sodium lines, which, by the way, on these spectra are not enhanced as compared with the nightglow intensity. The O*I* multiplet intensities seem, however, to vary considerably. This is especially apparent on the eight spectra taken from the 6250–6650 Å region. The intensity of the O*I* multiplet at 6455 Å varies quite differently from that of the circumjacent heads of the 8,5 first positive N_2 band. Neither does it seem to follow the variation of the forbidden O*I* lines. There is, on the other hand, an indication of a covariance of the O*I* multiplet with the N*II* line at 6482 Å and the H_α line.

The N*I* forbidden doublet at 5200 Å is relatively strong, but no other N*I* lines are observed between 5200 Å and 6650 Å.

The N*II* multiplet at 5660 Å to 5730 Å appears with six components of the correct intensity, and the singlet line at 6482 Å is present on all spectra of the red region, sometimes strong. The forbidden ($^1D_2$–$^1S_0$) line at 5755 Å is also, though weak, definitely present. The same applies to the two ($^3P_{1\cdot2}$–$^1D_2$) lines at 6548 Å and 6583 Å when they are not masked by N_2 first positive bands.

The H_α line is absent on three of the eight spectrograms taken of the region 6250 Å to 6650 Å, and strong on only one. This spectrogram was taken from approximately magnetic horizon and the line is symmetrically broadened with a halfwidth of about 15 Å.

NIGHTGLOW SPECTRA

From January to April 1958 were taken eleven successful nightglow spectrograms, three from the region 5800 Å to 6350 Å, the rest covering the range from 6200 Å to 6650 Å. No auroral activity was visible on these nights.

The most remarkable results obtained from these plates is the observation of a sharp H_α line and the presence of the forbidden N*II* ($^3P_{1\cdot2}$–$^1D_2$) lines.

The H_α line is visible on all spectra from the red region. Usually it is fainter than the $P_1(2\text{–}3)$ lines of the OH 6,1 band, but in three cases its intensity is from one to four times that of the *P*-lines. There is no observable Doppler broadening of the line.

In seven out of eight nights the stronger component (6583 Å) of the N*II* forbidden "doublet" was observed, the weaker one (6548) being present only on three spectrograms.

It seems to be a very close correlation between the intensities of the H_α line and the N*II* forbidden lines. When the former are strong, so are the latter. There is also a close correspondence between the intensity variation of the radiations along the spectral lines. In a couple of cases there appears more intense patches on all three lines at the same distance from the line ends, showing a simultaneous enhancement of the radiations in a comparatively restricted area of the sky.

Four bands of the OH rotation–vibration $X^2\Pi$ system (MEINEL, 1950a), viz. the

8,2, 5,0, 9,3,and 6,1, appear on the plates, the last two with an intensity suitable for temperature measurements.

Table 1 contains mean values from all plates of wavelengths and intensities of the rotational lines of the four bands.

Table 1. Wavelengths (λ) and intensities (I) of OH-bands

Band:	8,2		5,0		9,3		6,1	
Rot. lines*	λ	I	λ	I	λ	I	λ	I
R_1 (2,3)							6465·7	5·6
R_1 (1)	5866·9	6·0	6143·8	2·9	6235·7	12·2	71·0	5·0
R_2 (1)	74·5		51·0		42·1	3·6	78·3	2·1
Q_1 (1)	*M*:NaI†	(59)	70·4	3·1	58·2	14	6498·7	10
Q_1 (2)	*M*:NaI†	(32)			65·4	3·9	6505·1	4·2
Q_1 (3)							14·0	1·8
P_2 (2)	5906·8	2·4	6194·0		76·8	3·8	22·2	3·0
P_1 (2)	15·3	4·3	6202·7	2·2	6287·8	9·5	32·8	7·3
P_2 (3)	25·3	3·0			*M*:OI†	(170)	44·0	4·0
P_1 (3)	33·2	5·6	21·5	2·6	6307·4	10·8	53·7	8·0
P_2 (4)	45·3	2·1			21·7	4·6	69·1	3·2
P_1 (4)	53·8	4·0	42·1		30·4	8·9	77·3	6·3
P_2 (5)	70·7	1·5			49·0	3·2	6596·3	2·0
P_1 (5)	77·9	2·6	6268·4		56·8	5·6	6603·8	3·3
P_2 (6)	5995·7				79·4	1·8	27·6	0·9
P_1 (6)	6004·6				6386·8	3·1	6633·5	1·5
P_2 (7)					6414·1	1·1		
P_1 (7)					20·7	1·6		
P_2 (8)					51·5			
P_1 (8)					6457·6			
Sum intensity:		60		25		125		80
M quanta/cm²-sec		75		30		160		100

* Terminology as in HERZBERG (1951). The quantum number K refers to the lower level.
† *M:* = Masked by

The wavelength values are systematically smaller, by 1–5 Å than those given by SMALL and PETRIE (1952) and CABANNES *et al.* (1950), conforming somewhat better with CHAMBERLAIN and OLIVER's (1953) values.

The band lines of the 5,0 and 6,1 bands agree, however, within a fraction of an Ångström unit with values computed from energy levels of the OH $X^2\Pi$ state, given by DIEKE and CROSSWHITE (1948) ($v = 0 - 3$) and HERMAN and HORNBECK (1953) ($v = 4 - 6$). It should therefore be reasonably safe to conclude that the values given here of the 8,2 and 9,3 band lines should be correct to within at least 1 Å.

The intensities are measured from the photographic densities of the lines,

VOL. 16 1959

corrected for plate sensitivity and absorption in the spectrograph. Since the spectrograph does not cover the range of the four bands on one spectrogram, it has been necessary to use the 9,3 band, of which the two ranges have some lines in common, to bridge the gap. This may have introduced a small uncertainty in the intensities of the 5,0 and 8,2 bands relative to that of the 6,1 and 9,3 bands.

The sum intensities in the next but last row of the table are extrapolated from the line intensities on the assumption that the rotational temperature is 215°K (see later). Comparison of the sum intensities of the 8,2 and 9,3 bands with the intensities of the Na*I* and O*I* (6300) lines respectively, and use of the values given by ROACH (1956) of the absolute intensities (in R = M quanta/cm²-sec) of these lines has yielded the values in the last row of the table (means of the sets of numbers computed from the relative intensities).

The absolute intensities of the OH bands are probably rather uncertain. The intensity ratio of the 8,2 band and the Na*I* lines is fairly constant on all spectra (also those from weak aurorae), whereas the 9,3/O*I* (6300) intensity ratio may vary by a factor greater than 10, probably owing to invisible auroral activity. The intensity of the O*I*-line, given in the table, is the mean from two spectrograms exposed towards SW and which gave the lowest intensity of the line.

BARBIER (1955) has tentatively stipulated the absolute intensity of the OH 9,3 band to 70 R. In view of the uncertainties mentioned, the agreement is satisfactory.

It is usually assumed that the intensity of the OH 8,2 band is weak as compared with the intensity of the sodium doublet (see e.g. BARBIER, 1955). The sum intensity of the former is here found to be about one half of that of the latter. It is further found that in quiet nights the intensities of the 9,3 band and the O*I* (6300) line are about equal. If these results are correct, the correlation reported between the sodium doublet and the red O*I* lines in the nightglow (BARBIER, 1955; ROACH, 1956), found with photometric instruments using optical filters, may, at least partly, be due to the "background" of OH bands. The same applies to the correlation between the near infra-red OH band and the red O*I* line (BARBIER, 1955).

TEMPERATURE MEASUREMENTS

Eight nightglow spectrograms had a suitable density for measuring the rotational temperature of the 6,1 band from the lines in the *P*-branches. The 9,3 band was measurable on only four plates.

Assuming thermal equilibrium, the intensity, I, of a rotational line within a band is given by

$$I(J) = C\nu_J^4 \, s_J \exp\{-F(J)hc/kT\}$$

where ν_J is the wavenumber of the line, s_J the line strength and $F(J)$ the rotational term value of the upper level with rotational quantum number J. Usually ν_J is regarded as a constant, and $F(J)$ is approximated by $BJ(J+1)$, where B is the rotational constant. One procedure is then to plot $\ln(I/s)$ against $J(J+1)$. The rotational temperature T is then determined by the slope of the straight line obtained (if thermal equilibrium exists).

The lines in the *P*-branches of the OH bands are spaced to such an extent that

the variation of v_J should not be ignored. The variation has been taken into account here.

The OH bands are doublets, and the sub-branch, originating from the transition between $^2\Pi_{3/2}$ levels, the P_1-lines, has a rotational constant B_1 different from the rotational constant B_2 of the sub-branch P_2 (from $^2\Pi_{1/2}$ transitions). B_1 and B_2 are not given in the literature. Instead of using approximate values (HERZBERG, 1951, p. 233), and also to take into account second and higher order terms in $F(J)$,

Table 2. Rotational temperature (°K) in the OH-stratum of the atmosphere

		Upper		Lower		
Band	Number of spectra	P_1	P_2	P_1	P_2	Mean
6,1	8	216	217	212	213	214·5
9,3	4	214	216	214	215	215
Mean		216		213·5		215

the experimental energy values of the rotational levels have been used directly. (Values for $v = 6$ were taken from HERMAN and HORNBECK (1953). The 9,3 band lines in Table 1, combined with DIEKE and CROSSWHITE's (1948) term values for $v = 3$, supplied the $v = 9$ terms.)

The expressions for s_J are from HÖNL and LONDON (1925):

$$s_J = \begin{cases} \dfrac{(J+1)^2 - 9/4}{J+1}, & P_1 \\ \dfrac{(J+1)^2 - 1/4}{J+1}, & P_2 \end{cases}$$

In the experimental arrangement, the slit of the spectrograph was placed vertically. The two end parts of the spectral lines, corresponding to the upper and lower part of the observed area of the sky, were measured. Each spectrum so provided four independent temperature measurements, one from each sub-branch on both end parts.

Table 2 contains mean values of the temperature measured from the number of spectra quoted in the second column. The range of the individual values, on which the means are based, is from 205°K to 221°K.

The temperature found is considerably lower than that given by MEINEL* (1950b) and KRASSOWSKY *et al.* (1958).

* The line intensities of the 5,1 and 6,2 bands given by MEINEL, treated in the manner adopted here, give temperatures of 251 and 247°K instead of his values 265 and 258°K.

VOL. 16 1959

The last named authors report temperatures ranging from 230°K to 330°K, dependent on the direction of observation. No such effect has been observed here although the direction of observation has been varied from SW to NE.

The good correspondence of the temperatures measured from the two sub-*P*-branches and the two different bands, and the small spread in the individual determinations, should indicate that the temperatures in Table 2 are tolerably accurate. If this is correct, the height of the atmospheric stratum, emitting the OH bands, must, according to the temperature curve of the Rocket Panel (1952), lie either just above the 70 km level or just below 100 km (alternatively 90 km).

Additional evidence is required to settle the question: which of the two heights is the correct one? This may be supplied by the observation mentioned above that the intensities of the OH bands seem not to be enhanced during auroral displays as compared with the intensities on quiet nights. According to KRASSOWSKY (1957) the formation of the OH molecules takes place near the 100 km level by a reaction between excited oxygen molecules and hydrogen atoms, the excited oxygen molecules being formed from atomic oxygen. It is difficult to understand how the OH formation could remain approximately the same during quiet nights as during aurorae when in the latter case the available energy should be more than sufficient to increase the oxygen dissociation, and the incoming protons should supply a greater concentration of hydrogen atoms. The theory of Bates and Nicolet (BATES *et al.*, 1956), by which the OH molecules are formed by a reaction between ozone and atomic hydrogen around the 70 km level, is not influenced by this argumentation. The author is therefore of the opinion that the results here given are in favour of the ozone–hydrogen hypothesis of Bates and Nicolet, and that the emission of the OH bands takes place at a height of about 70 km above the surface of the earth.

Acknowledgements—The author is very much indebted to the Norwegian Research Council for Science and the Humanities for financial support of this investigation. He also wishes to express his thanks to cand. real Öivind Haugerud for valuable assistance with the observations and computations.

REFERENCES

BARBIER D.	1955	*The Airglow and the Aurorae.* Pergamon Press, London.
BATES D. R. and MOISEIWITSCH B. L.	1956	*J. Atmosph. Terr. Phys.* **8,** 305.
CABANNES J., DUFAY J. and DUFAY M.	1950	*C.R. Acad. Sci., Paris* **230,** 1233.
CHAMBERLAIN J. W. and OLIVER N. J.	1953	*Phys. Rev.* **90,** 1118.
DIEKE G. H. and CROSSWHITE H. M.	1948	Bumblebee Series, Report No. 87. Johns Hopkins University.
HERMAN R. C. and HORNBECK G. A.	1953	*Astrophys. J.* **118,** 214.
HERZBERG G.	1951	*Spectra of Diatomic Molecules.* Van Nostrand, New York.
HÖNL H. and LONDON F.	1925	*Z. Phys.* **33,** 803.
KRASSOWSKY V. I.	1957	*J. Atmosph. Terr. Phys.* 10, 49.
KRASSOWSKY V. I., GALPERIN G. I. MIRINOW A. V., PROKUDINA V. S., SHEFOV N. N., FEDOROVA N. I. and BAGARIAZKY B. A.	1958	*Ann. Géophys.* **14,** 356.

MEINEL A. B.	1950a	*Astrophys. J.* **111,** 555.
MEINEL A. B.	1950b	*Astrophys. J.* **112,** 120.
PETRIE W. and SMALL R.	1952	Scientific Rep. No. AR–7. University of Saskatchewan, Saskatoon.
ROACH F. E.	1956	*I.A.G.A. Bul.* No. 15b, 100.
SMALL R. and PETRIE W.	1952	Scientific Rep. No. AR–8. University of Saskatchewan, Saskatoon.
The Rocket Panel	1952	*Phys. Rev.* **88,** 1027.

VOL.
16
1959

Journal of Atmospheric and Terrestrial Physics 1959, Vol. 16. pp. 259 to 262. Pergamon Press Ltd. Printed in Northern Ireland

The fluctuations of the acceleration of satellites

H. K. PAETZOLD
Max-Planck-Institute of Aeronomy, Institute of Physics of the Stratosphere,
Weisenan, Kreis Ravensburg, Germany

(*Received* 5 *January* 1959)

OWING to the air drag the satellites are accelerated continuously. The decrease dP/dn of the revolution time P can be given with good approximation by:

$$\frac{dP}{dn} = -3P\frac{\bar{F}}{M}\oint_{\text{orbit}}\rho(h)\,ds, \tag{1}$$

Where $\bar{F}$ is the mean effective cross-section, M the mass of the satellite and $\rho(h)$ the air density at the height h. The integral represents the air mass, which is hit by the satellite during one revolution (PAETZOLD, 1958a).

Equation (2) can be written in the modified form:

$$\frac{dP}{dn} = -3P\frac{\bar{F}}{M}\rho(h_{Pe})2\pi(R + h_e)c(e) \tag{2}$$

where h_{Pe} is the perigee's height and R the earth's radius. The factor $c(e)$ gives the influence of the eccentricity of the satellite's orbit. In the possible boundaries $c(e)$ does not vary very strongly with the scale height of the atmosphere. It is well known, that from the observed accelerations of the different satellites a much higher density of the upper atmosphere results than given by the ARDC model 1956 (PAETZOLD, 1958b; PAETZOLD and ZSCHÖRNER, 1958b; SCHILLING and STERNE, 1958; WITNEY, 1958b; SCHILLING and WITNEY, 1958). Further, the density seems to decrease more slowly with height in the higher latitudes than in the lower. From this it is probable, that the upper atmosphere is heated up considerably by the influx of charged particles. One part may come directly from the sun, the other from the interplanetary hydrogen, which will be totally ionized. The density of the interplanetary hydrogen is about 500 protons/cm^3 (ELSÄSSER, 1953).

This general picture concerns only the mean conditions. However, considerable variations of the accelerations with time have been observed (JACCHIA and BRIGGS, 1958; KING-HELE, 1958a, b; HERGENHAHN, 1958). The explanation is somewhat complicated owing to the fact that satellites are not spherical in shape with the exception of $1957\alpha_2$ and $1958\beta_2$. From the fadings of the radio signals it is probable that some satellites show a precession besides the spinning (PAETZOLD and ZSCHÖRNER, 1958a). By this the variation of the cross-section $\bar{F}$ may show regular periods in time, as was found by WITNEY (1958a). But the main part of the fluctuations of the acceleration seems to be caused by the atmosphere itself.

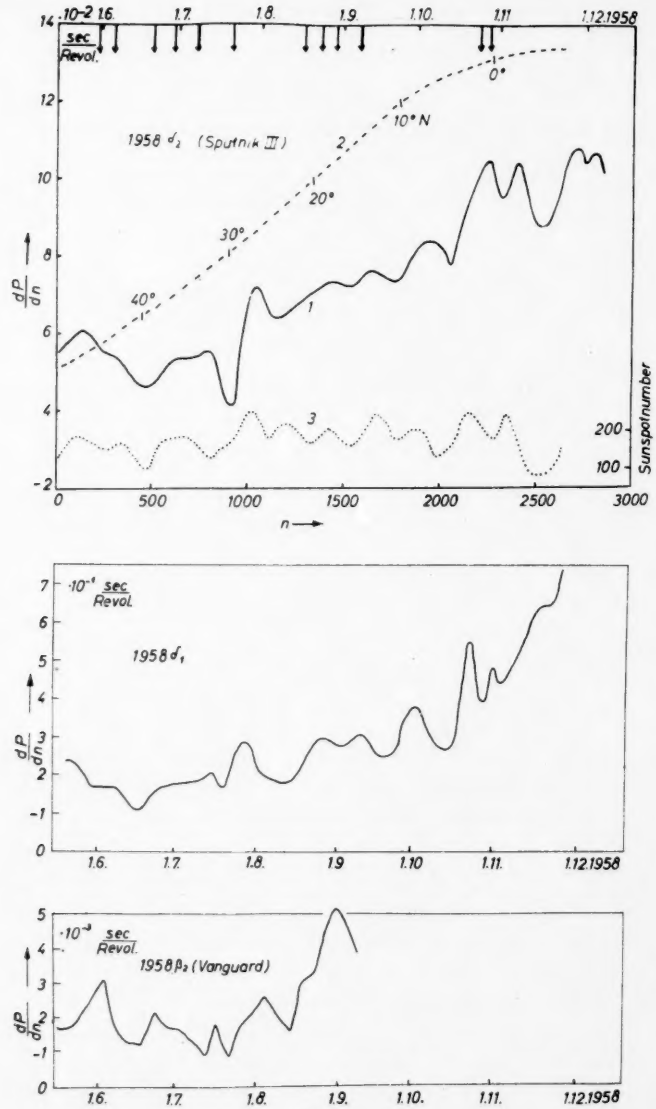


Fig. 1.

Fig. 1 shows the irregularities of *Sputnik* III (1958 δ_2) as they have been observed (PAETZOLD, 1959) with a direction finder at 20 Mc/s.* The accuracy in time was about 2 sec, while the irregularities of the satellite often amounted to several minutes. The revolution of the rocket of *Sputnik* III (1958 δ_2) seems to undergo the same fluctuations after the optical observations in spite of the fact that the two satellites have very different shapes. Further, the acceleration of *Vanguard* (1958 β_2) also shows the pronounced decrease during the first half of June 1958,

* The direction finder has been developed by the firm Telefunken, Ulm.

VOL. 16 1959

and the following increase at the end of July, as the Russian satellite did. It is striking that the spherical *Vanguard* shows much stronger irregularities than the rockets of *Sputnik* II *and Sputnik* III.

In Table 1 the mean relative amplitude of the fluctuations are given for four satellites. It does not seem reasonable that the effect of electrical forces on the satellites will vary to such a degree. Therefore variations of the air density seem to be most probable. They will take place predominantly in the vicinity of the perigee, because the factor $c(e)$ in equation (2) depends very little upon the scale

Table 1

Satellite	1957β	1958_1	1958_2	$1958\beta_2$
Initial height of perigee (km)	225	226	226	656
Variation of dP/dn	15%	30%	20%	70%

height. Obviously the amplitude of the variations increases with height, at 1000 km they may amount to a factor of 10. These fluctuations seem to be mainly temporary and not local, as it has been supposed already (KING-HELE and WALKER, 1958), for the period for the Russian and the American satellites is equal in spite of the very fast shift of the perigee of *Vanguard*. In the mean, every 10–15 days a significant change of the acceleration occurs. The fluctuations seem to be predominantly irregular. No direct correlations could be found with special events such as magnetic and ionospheric storms. But there may be a correlation to the sunspots number. If this had been found valid, the density and the temperature would show a pronounced variation during the solar cycle. In any case the fluctuations of the density of the upper atmosphere will probably be caused by a variable heating (BARTELS, 1958). If no correlation existed with the solar activity, the density of the interplanetary protons would vary. In this case clouds of hydrogen would exist with a size of 30–60 $\times$ 10^6 km dia.

REFERENCES

BARTELS J.	1958	*Naturwissenschaften.* **45,** 181.
ELSÄSSER H.	1953	*Z. Astrophys.* **33,** 274.
HERGENHAHN G.	1958	*Naturwissenschaften.* **45,** 434.
JACCHIA L. G. and BRIGGS R. E.	1958	*Smithsonian Inst. Spec. Rep.* No. 18, 9.
KING-HELE D. G.	1958a	*Nature, Lond.* **181,** 172.
KING-HELE D. G.	1958b	*Nature, Lond.* **182,** 1409.
KING-HELE D. G. and WALKER D. M. C.	1958	*Nature, Lond.* **182,** 860.
PAETZOLD H. K.	1958a	*Raketentechnik u. Raumfahrtforschg.* **2,** 50.
PAETZOLD H. K.	1958b	*Umschau* **58,** 1.

PAETZOLD H. K.	1959	*Planetary Space Sci.* **1,** 115.
PAETZOLD H. K. and ZSCHÖRNER H.	1958a	*Telefunken-Zeitschr.* **31,** 100.
PAETZOLD H. K. and ZSCHÖRNER H.	1958b	*Naturwissenschaften.* **45,** 485.
SCHILLING G. F. and STERNE T. E.	1958	*IGY Satellite Rep. No.* 430.
SCHILLING G. F. and WHITNEY C. A.	1958	*Smithsonian Inst. Spec. Rep.* No. 18, 13.
WHITNEY C. A.	1958a	*IGY Satellite Rep.* No. 3, 6.
WHITNEY C. A.	1958b	*IGY Satellite Rep. No.* 422.

VOL.
16
1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16. pp. 263 to 273. Pergamon Press Ltd. Printed in Northern Ireland

Note on the tidal theory of the S_q magnetic field*

R. L. INGRAHAM

New Mexico College of Agriculture and Mechanic Arts, State College, New Mexico

(*Received 5 January* 1959)

Abstract—The observed ratio of the diurnal to the semidiurnal components of the S_q field is greater than unity, although one would expect a ratio about a thousand times smaller on order of magnitude tidal theory. We therefore have made an exact calculation using a dynamo theory, with a certain analytically manageable model of diurnal and semidiurnal tides derived from the linearized Taylor-Pekeris theory to see whether "accidental" features might arise. It is shown that the anomaly persists. Possible explanations of the puzzle based on the details of this tidal theory are suggested. (1) There may be a resonating diurnal mode besides the well-established semidiurnal one. (2) Diurnal and semidiurnal motions rise to high maxima around latitudes 30°N and S and in the polar regions, respectively, as a result of the rotation of the atmosphere. This feature is ignored in both order of magnitude theory and in our model. The diurnal maxima should have much greater weight than the semidiurnal in the integral for the S_q field.

1. INTRODUCTION

ONE of the puzzles associated with the solar quiet day (S_q) magnetic field is that the diurnal component is roughly a thousand times as great as one would expect from order of magnitude arguments based on the tidal theory of atmospheric oscillations. In a word, dynamo theory requires that the diurnal and semidiurnal S_q amplitudes at the ground be in the same ratio as the corresponding characteristic tidal velocities at 100 km, the probable seat of dynamo action. The barometric data and atmospheric tidal theory (TAYLOR, 1936; PEKERIS, 1937), suggest a value of about 1/500 for this ratio. But experimentally this ratio is found to be somewhat greater than unity. The details of this order of magnitude argument are given in the following section.

2. DETAILS OF THE ORDER OF MAGNITUDE CALCULATION

Let $p_1(r)$ and $p_2(r)$ be the amplitudes of the tidal pressure oscillations† at distance r measured from the earth's centre, reckoned for convenience relative to a spherical harmonic latitude and longitude dependence:‡‡

$$\begin{aligned} p_1(r, \theta, \phi) &\equiv p_1(r)\, Y_1^1(\theta, \phi) e^{i\alpha_1} \\ p_2(r, \theta, \phi) &\equiv p_2(r)\, Y_2^2(\theta, \phi) e^{i\alpha_2}. \end{aligned} \tag{2.1}$$

That is, assume tidal gravitational potentials $\Omega_1 \propto Y_1^1$, $\Omega_2 \propto Y_2^2$. The magnetic data support this assumption well; the barometric data, not so well. For instance, it is difficult to find a single simple harmonic for the latitude dependence of p_1; whereas for p_2 one finds, for the part depending on local time, that $\sin^3 \theta$ seemingly

* This work was supported by the Office of Naval Research.

† The subscripts 1 and 2 will always refer to diurnal (frequency $= 1\omega$) and semidiurnal (frequency $= 2\omega$) quantities, where ω is the rotational frequency of the earth.

‡‡ Here θ and ϕ are colatitude and azimuth reckoned in a non-rotating frame. Thus $\phi \equiv$ local solar time measured in degrees. The solar tidal motions involve the longitude ϕ_r and Greenwich time t to good approximation only in the combination $\phi \equiv \phi_r + \omega t$. Throughout this article, the convention on complex quantities will be: Take Re. Note also that we use the definition of the Y_l^m used by ROSE, (1955), CONDON and SHORTLEY, (1953) and others; N.B. their sign conventions.

expresses the latitude dependence better than $Y_2^2(\propto \sin^2\theta)$. The empirical formulae are (MITRA, 1952)

$$\begin{aligned} p_1 &= 0{\cdot}30 \sin(\phi + 0) \text{ mm Hg (near the equator)} \\ (p_2)_{\text{travelling wave}} &= 0{\cdot}937 \sin^3\theta \sin(2\phi + 154^\circ) \text{ mm Hg.} \end{aligned} \tag{2.2}$$

From these we infer $p_1(a) = 1{\cdot}15 \times 10^3$ dyn-cm^{-2}, $p_2(a) = 3{\cdot}22 \times 10^3$ dyn-cm^{-2} for the amplitudes in (2.1) at the earth's surface, $r = a$. With these numbers construct the characteristic velocity amplitudes $V_i(r)$ $(i = 1, 2)$ occurring in the tidal theory:

$$V_i(r) \equiv \frac{1}{r\omega}\left(\frac{p_i(r)}{\rho_0(r)} + \Omega_i\right) \cong \frac{p_i(r)}{\rho_0(r) r\omega} \tag{2.3}$$

where $\rho_0(r)$ is the undisturbed atmospheric density at r. At sea level, $V_1(a)$, $V_2(a) \cong 10^{-4}$ km-sec^{-1} and $V_2(a)/V_1(a) \cong 3$. Now according to the PEKERIS theory, the semidiurnal mode resonates, which means both that a small Ω_2 can induce uniformly large oscillations at all heights, and that there is *magnification*, that is, the amplitudes increase rapidly with height. One finds* that $V_2(a + 100) = 170\ V_2(a)$, where a will always be assumed measured in kilometres. So far no one has found a resonating diurnal mode from the PEKERIS theory. Therefore one *expects* no appreciable magnification: $V_1(a + 100) \cong V_1(a)$. Combining these equations,

$$V_2(a + 100) \cong 510 V_1(a + 100). \tag{2.4}$$

Here the only result that we have actually required from the detailed tidal theory is the value 170 of the semidiurnal magnification.

Now for the magnetic field at the ground due to the dynamo action of these tidal motions across the earth's dipole field $\mathbf{B}_0$ in the thin ionized layer of conductivity κ about 100 km high, from Ampère's law (in heaviside-lorentz units)

$$\mathbf{B}(\mathbf{r}) = \frac{1}{4\pi c}\int_{\text{ionized layer}} \frac{\nabla' \times \left[\frac{\kappa}{c}\left(\mathbf{v} \times \mathbf{B}_0\right) + \mathbf{j}_{\text{ret}}\right]}{R}\, d^3x' \tag{2.5}$$

where $R \equiv |\mathbf{r} - \mathbf{r}'|$ and we shall be interested in $r = a$. The "return" current $\mathbf{j}_{\text{ret}}$ is an irrotational current density† which guarantees $\nabla \cdot \mathbf{j}_{\text{total}} = 0$; assume strict electrical neutrality $\sigma \equiv 0$. Now where $\mathbf{B} = -\nabla\Omega$, factoring dimensional factors out of the integral, then

$$U_i(a, \theta, \phi) = a\, K_m V_i(a + 100) \frac{M}{(a + 100)^3} \Phi_i(\theta, \phi) \qquad (i = 1, 2) \tag{2.6}$$

where $K_m \equiv \int \kappa_m\, dr$ is the "total" conductivity of the layer in e.m.u. ($\kappa_m \equiv \kappa/4\pi c^2$); for $\mathbf{v}_i$ a tidal velocity, V_i is the characteristic velocity amplitude defined by (2.3); and all the pure number factors resulting from the integration are lumped into the θ,ϕ dependence Φ_i. Experimentally, one finds that the angular dependences are mainly Y_2^1 and Y_3^2 for $i = 1$ and 2, respectively. So if

$$\Phi_1 \equiv \varepsilon_1 Y_2^1 e^{i\alpha_1}, \qquad \Phi_2 \equiv \varepsilon_2 Y_3^2 e^{i\alpha_2} \tag{2.7}$$

* According to our numerical calculations. There is also a sign reversal, which does not concern us here.

† Cf. CHAPMAN's discussion (CHAPMAN and BARTELS, 1951, p. 769). The gradient of the ion pressure also contributes in general to this return current (cf. SPITZER, 1956, equation (2–22)). Of course $\mathbf{j}_{ret}$ contributes nothing to the resultant magnetic field.

VOL. 16 1959

where the ε_i are dimensionless (positive) numbers, the ratio of the magnetic potential amplitudes $U_i(r)$ at $r = a$ is simply

$$\frac{U_2(a)}{U_1(a)} = \frac{\varepsilon_2 V_2(a + 100)}{\varepsilon_1 V_1(a + 100)}. \tag{2.8}$$

The empirical formula due to CHAPMAN and BARTELS (1951, p. 692), manipulated a bit to express the angular parts in terms of spherical harmonics, is at $r = a$

$$U = U_1 + U_2 = a[12 Y_2^1 e^{i210^\circ} + 6{\cdot}3 Y_3^2 e^{i210^\circ}]\gamma \text{ km}. \tag{2.9}$$

The units are γ km, where $1\gamma = 10^{-5}$ G. This U, it should be mentioned, includes only the diurnal and semidiurnal modes; it is also restricted to that part of the S_q potential due to sources external to the earth; finally, we have chosen this potential for a typical sunspot minimum year in order to minimize the contribution of extraneous, nontidal phenomena. Hence the empirical amplitude ratio is

$$\left(\frac{U_2(a)}{U_1(a)}\right)_{\text{emp}} \cong \frac{1}{2}. \tag{2.10}$$

Now an important assumption: one expects integrals of products of spherical harmonics over the unit sphere not to yield abnormally large or small numbers. That is, one *expects* $\varepsilon_1 \cong \varepsilon_2 = 0(1)$. But then (2.8) with (2.4) gives the theoretical value

$$\frac{U_2(a)}{U_1(a)} \cong 510. \tag{2.11}$$

Comparing this with (2.10), we see that there is a discrepancy of about a factor 10^3 between experiment and order of magnitude tidal theory of S_q.

3. Quantitative Treatment

In view of this discrepancy, it is useful to evaluate the integral (2.5) exactly, using for the tidal velocities those given by the TAYLOR–PEKERIS theory (modified in a certain way to eliminate unphysical divergences). The most obvious use of such a calculation is to check the validity of the assumption that the pure numbers ε_1 and ε_2 are of the same order of magnitude. Indeed, if one of the integrals should show an "accidental" feature such that $\varepsilon_1 \cong 10^3\, \varepsilon_2$, the discrepancy would vanish. A further service a precise calculation might perform is to see what, if any, higher harmonics are predicted in the latitude–longitude dependences (2.7).

For the polar coordinate velocity components u, v, w PEKERIS' formulae (9) and (10) are used. The parameters are $\sigma_1 = \omega$, $s_1 = 1$, $f_1 = \frac{1}{2}$ and $\sigma_2 = 2\omega$, $s_2 = 2, f_2 = 1$ for the diurnal and semidiurnal modes, respectively. The horizontal components u and v are defined in terms of $Q \equiv p/\rho_0 + \Omega$ and for the two modes we write

$$\begin{aligned} Q_1(r, \theta, \phi) &\equiv Q_1(r)\, Y_1^1(\theta, \phi), \qquad w_1(r, \theta, \phi) \equiv w_1(r)\, Y_1^1(\theta, \phi) \\ Q_2(r, \theta, \phi) &\equiv Q_2(r)\, Y_2^2(\theta, \phi), \quad w_2(r, \theta, \phi) \equiv w_2(r)\, Y_2^2(\theta, \phi). \end{aligned} \tag{3.1}$$

Incidentally ϕ here is *not* PEKERIS' coordinate ϕ (the longitude); θ and ϕ are as defined in footnote on p. 263. We thus approximate the exact angular eigenfunctions by their first terms Y_1^1 and Y_2^2 (in addition to changing the normalization used by PEKERIS), which is permissible (see WILKES, 1949, pp. 43, 44). This semidiurnal mode is in fact the resonating one with equivalent-depth eigenvalue $H = 7{\cdot}87$ km, discussed in detail by PEKERIS; its exact angular eigenfunction ψ is given by the first line of his formula (15). Although the diurnal mode here was not discussed by PEKERIS, we assume that it exists since the magnetic data point to such a mode. Certain parameters of this mode are consequently unknown, in particular its equivalent depth. It turns out, however, that this information can be dispensed with.

Owing to the assumption of a rotating atmosphere in the PEKERIS theory, the semidiurnal horizontal velocities have the factor $(f_2^2 - \cos^2\theta)^{-1} = \sin^{-2}\theta$. This leads to the defect that u_2 and v_2 not only do not satisfy the boundary conditions that they should vanish at the poles, but they actually diverge there. This unphysical feature seems to have been somewhat tactfully avoided in former treatments of this theory. Similarly, for any diurnal mode, u_1 and v_1 have the factor $(f_1^2 - \cos^2\theta)^{-1} = (\frac{1}{4} - \cos^2\theta)^{-1}$ which leads to infinite horizontal velocities at the latitudes 30°N and S. What this means physically, of course, is that the actual horizontal velocities rise to high but finite maxima, and hence the linearized theory breaks down, in the neighbourhood of these latitudes. Of this, more will come later, in Section 5. We have chosen to cope with this inadequacy of the linearized theory by using, instead of the singular PEKERIS velocities, the following well-behaved "cut-off" velocities

$$\begin{aligned}(u_{1\text{cutoff}}, v_{1\text{cutoff}}) &\equiv (\tfrac{1}{4} - \cos^2\theta)(u_1, v_1)\\ (u_{2\text{cutoff}}, v_{2\text{cutoff}}) &\equiv \sin^2\theta(u_2, v_2).\end{aligned} \tag{3.2}$$

(N.B. This replacement is not exactly equivalent to treating a non-rotating atmosphere. It is true that for a non-rotating atmosphere, $\omega \to 0$, the offending factors mentioned above are not present; but also the second term in u and the first term in v in PEKERIS' formula (10) drop out. Thus u and v for the non-rotating atmosphere form an irrotational horizontal flow, but, as can be verified, u_{cutoff} and v_{cutoff} do not.)

For the earth's dipole field $\mathbf{B}_0$, neglecting the tilt of the geomagnetic axis, write $-M/r^3(2\cos\theta, \sin\theta, 0)$ for its r, θ and ϕ components, respectively. Thus $-M/r^3$ ($M > 0$) is the equatorial value of the dipole field at height r.

Then the integral in Ampère's law (2.5) has been evaluated exactly, using the velocities (3.2) and this approximate geomagnetic field. The conductivity κ is allowed to be any function of r vanishing outside the shell $A \leq r \leq B$. A convenient formula* permits the magnetic potential at any point exterior to the shell to be written down once the components of $\mathbf{j}$ in some one fixed set of inertial axes

* Proof to be published elsewhere. Cf. the similar in spirit derivation of CHAPMAN (CHAPMAN and BARTELS, 1951) §23.12. CHAPMAN neglected vertical tidal velocity and required an irrotational horizontal flow. Our formula is more general in allowing any tidal velocities whatever, provided only that the current density be expressible as a sum of terms like (3.3).

VOL. 16 1959

are known. We have found that the most convenient set of components for this purpose are the circular cartesian components $j_\sigma(\sigma = +1, -1, 0)$: $j_+ \equiv j_x + ij_y$, $j_- \equiv j_x - ij_y$, $j_0 \equiv j_z$. Assume

$$(j_l{}^m)_\sigma = (f_l{}^m)_\sigma Y_l{}^{m+\sigma}(\theta, \phi) \tag{3.3}$$

which is the case in general and is true here in particular. The $(f_l{}^m)_\sigma$ are, in general, r-dependent. By linearity, one can treat each harmonic (l, m) separately, then add the results. The convention on complex quantities is always: Take Re. Now since $\nabla \times \mathbf{j}_{\text{ret}} = 0$, this term does not contribute to $\mathbf{B}$; so let $\mathbf{j}$ henceforth stand for the dynamo current density $\kappa/c(\mathbf{v}_{\text{cutoff}} \times \mathbf{B}_0)$ alone. The semidiurnal $\mathbf{j}$ is the sum* of three harmonics with $(l, m) = (1, 2)$, $(3, 2)$ and $(5, 2)$. The corresponding amplitudes $\mathbf{f}_l{}^m$ are

$$\begin{aligned}
\mathbf{f}_1{}^2 &= \frac{\kappa}{c}\frac{M}{r^3} V_2\left(0, \quad -\sqrt{\frac{1}{5}}\left(\frac{8}{7} + 2i\frac{w_2}{V_2}\right), \quad 0\right) \\
\mathbf{f}_3{}^2 &= \frac{\kappa}{c}\frac{M}{r^3} V_2\left(\sqrt{\frac{6}{7}}\left(\frac{4}{3} + i\frac{w_2}{V_2}\right), \quad -\sqrt{\frac{2}{35}}\left(8 - i\frac{w_2}{V_2}\right), \quad -4\sqrt{\frac{1}{7}}\right) \\
\mathbf{f}_5{}^2 &= \frac{\kappa}{c}\frac{M}{r^3} V_2\left(\frac{4}{21}\sqrt{\frac{42}{11}}, \quad -\frac{4}{7}\sqrt{\frac{1}{11}}, \quad -\frac{2}{7}\sqrt{\frac{7}{11}}\right).
\end{aligned} \tag{3.4}$$

The order of the circular components f_σ will always be (f_+, f_-, f_0). Here V_2 and w_2 are the r-dependent amplitudes defined by (3.1) (note $V_i(r) \equiv Q_i(r)/r\omega$); κ also is a function of r. According to tidal theory iw/V is real. The radicals come, of course, from converting sines and cosines into spherical harmonics, and from combining these spherical harmonics by the addition theorem. For the diurnal $\mathbf{j}$, which is also the sum of three harmonics $(l, m) = (0, 1)$, $(2, 1)$ and $(4, 1)$, one finds* (writing $f_{\text{diurnal}} \equiv g$)

$$\begin{aligned}
\mathbf{g}_0{}^1 &= \frac{\kappa}{c}\frac{M}{r^3}\frac{V_1}{4}\left(0, \quad -\sqrt{\frac{2}{3}}\left(\frac{19}{5} + 4i\frac{w_1}{V_1}\right), \quad 0\right) \\
\mathbf{g}_2{}^1 &= \frac{\kappa}{c}\frac{M}{r^3}\frac{V_1}{4}\left(\sqrt{\frac{1}{5}}\left(\frac{12}{7} + 8i\frac{w_1}{V_1}\right), \quad -\sqrt{\frac{2}{15}}\left(\frac{85}{7} - 4i\frac{w_1}{V_1}\right), \quad -\frac{39}{35}\sqrt{5}\right) \\
\mathbf{g}_4{}^1 &= \frac{\kappa}{c}\frac{M}{r^3}\frac{V_1}{4}\left(\frac{24}{7}\sqrt{\frac{1}{15}}, \quad -\frac{8}{35}\sqrt{6}, \quad -\frac{24}{35}\sqrt{\frac{5}{6}}\right).
\end{aligned} \tag{3.5}$$

The magnetic potential due to the current (3.3) at a point inside the shell is

$$U_l{}^m = -i\frac{2\pi}{l(2l+1)} P_l{}^m r^l Y_l{}^m \quad (r \leq A;\ l > 0) \tag{3.6}$$

where

$$\begin{aligned}
P_l{}^m \equiv [\sqrt{\{(l-m)(l+m+1)\}}\,(R_l{}^m)_+ + \sqrt{\{(l+m)(l-m+1)\}}\,(R_l{}^m)_- \\
+ 2m(R_l{}^m)_0]
\end{aligned} \tag{3.7}$$

$$(R_l{}^m)_\sigma \equiv \frac{1}{4\pi c}\int_A^B dr\, r^{1-l}(f_l{}^m)_\sigma.$$

* Some details are given in the appendix.

For completeness, for $l = 0$, r^l/l should be replaced by $-\log(a/r)$ in (3.6), where this U is normalized to vanish at $r = a$. From (3.6) one can read off the important facts (1) U_l^m is a solution of Laplace's equation, as it must of course be; (2) a current harmonic (l, m) induces only the *same* harmonic in the potential (a fact which seems not at all obvious to us, in view of the large number of harmonics which a given current harmonic spawns in the integrand of (2.5) by the repeated application of the addition theorem); (3) a given non-zero current harmonic $\mathbf{j}_l^m$ need *not* have a resultant magnetic field, the necessary and sufficient condition that it have no field being: $P_l^m = 0$. The crucial quantities P_l^m for the semidiurnal mode come out

$$P_1^2 = P_5^2 = 0,\ P_3^2 = 8\sqrt{\frac{1}{7}}\left(-3 + i\,\frac{\bar{w}_2}{\bar{V}_2}\right)\frac{1}{\bar{r}^2}\,K_m\,\frac{M}{\bar{r}^3}\,\bar{V}_2 \tag{3.8}$$

and for the diurnal mode

$$P_0^1 = P_4^1 = 0,\ P_2^1 = 2\sqrt{\frac{1}{5}}\left(-4 + 3i\,\frac{\bar{\bar{w}}_1}{\bar{\bar{V}}_1}\right)\frac{1}{\bar{\bar{r}}}\,K_m\,\frac{M}{\bar{\bar{r}}^3}\,\bar{\bar{V}}_1. \tag{3.9}$$

Here we have introduced the total conductivity $K_m \equiv \int_A^B dr\kappa_m(r)$ in e.m.u., and the barred quantities are certain averages over the depth of the layer defined in an obvious way. The quantities P_1^2 and P_0^1 had to vanish because Y_1^2 and Y_0^1 are meaningless, but it seems remarkable that P_5^2 and P_4^1 vanish also! Hence the diurnal and semidiurnal magnetic potentials at $r = a$ come out to be

$$U_1 = (U_1)_2^1 = aK_m\bar{\bar{V}}_1\,\frac{M}{\bar{\bar{r}}_3}\,i\,\frac{2\pi}{5}\sqrt{\frac{1}{5}}\left(4 - 3i\frac{\bar{\bar{w}}_1}{\bar{\bar{V}}_1}\right)\left(\frac{a}{\bar{\bar{r}}}\right)Y_2^1(\theta, \phi) \tag{3.10}$$

$$U_2 = (U_2)_3^2 = aK_m\bar{V}_2\,\frac{M}{\bar{r}^3}\,i\,\frac{16\pi}{21}\sqrt{\frac{1}{7}}\left(3 - i\,\frac{\bar{w}_2}{\bar{V}_2}\right)\left(\frac{a}{\bar{r}}\right)^2 Y_3^2(\theta, \phi).$$

These must still be multiplied by phase factors $e^{i\beta_1}$, $e^{i\beta_2}$ to make the phases agree with the empirical formula (2.9); these phases were left undetermined in the tidal velocities.

4. Discussion of the Results

The first notable fact is that the use of the detailed form of the (cut off "linearized") tidal velocities in (2.5) does not produce any higher harmonics than are given by (2.9). These terms, which are small but exist, must then be due to the discrepancy between the actual tidal velocities and this approximation of them, or to some other cause.

We compute the pure numbers ε_i defined by (2.7) and (2.6):

$$\begin{aligned}\varepsilon_1 &= \frac{2\pi}{5}\sqrt{\frac{1}{5}}\left(4 - 3i\,\frac{\bar{\bar{w}}_1}{\bar{\bar{V}}_1}\right)\left(\frac{a}{\bar{\bar{r}}}\right) \cong \frac{8\pi}{5}\sqrt{\frac{1}{5}}\left(\frac{a}{a + 100}\right) = 2{\cdot}21 \\ \varepsilon_2 &= \frac{16\pi}{21}\sqrt{\frac{1}{7}}\left(3 - i\,\frac{\bar{w}_2}{\bar{V}_2}\right)\left(\frac{a}{\bar{r}}\right)^2 \cong \frac{16\pi}{7}\sqrt{\frac{1}{7}}\left(\frac{a}{a + 100}\right)^2 = 2{\cdot}63\end{aligned} \tag{4.1}$$

since $w(r)$ is usually considerably less than $V(r)$, and $a/(a + 100) = 6370/6470$

VOL. 16 1959

$= 0.985$. Hence there is no "accidental" feature for this model: the assumption $\varepsilon_1 = \varepsilon_2 = 0(1)$ of Section 2 is good. The ratio of the velocity amplitudes at 100 km as given by the quantitative theory is

$$\frac{V_2(a+100)}{V_1(a+100)} = \frac{\varepsilon_1}{\varepsilon_2}\left(\frac{U_2(a)}{U_1(a)}\right)_{\text{emp}} = \frac{2{\cdot}21}{2{\cdot}63} \times \frac{6{\cdot}3}{12} = 0{\cdot}44. \tag{4.2}$$

We can also get an estimate of the size of the total conductivity of the ionized region causing the S_q field by comparing, say, the more trustworthy semidiurnal parts of (3.10) and (2.9). For the velocity $\bar{V}_2$ at 100 km we prefer to use a semi-empirical determination $\bar{V}_2 = 170\, V_2(a)$ in which only the value 170 of ground-to-100 km magnification is required from the PEKERIS theory. $V_2(a)$ can be computed from the barometric data at sea level given in Section 2; we need also the data $\rho_0(a) = 1{\cdot}25 \times 10^{-3}$ g-cm^{-3}, $a = 6370$ km and $\omega = 7.29 \times 10^{-5}$ sec^{-1}. This gives $\bar{V}_2 = 9{\cdot}30 \times 10^3$ cm-sec^{-1} ($= 334$ km-hr^{-1}), a large velocity. Also the geomagnetic field on the equator at 100 km $M/\bar{r}^3 = (a/\bar{r})^3 M/a^3 = (0.985)^3\, 3{\cdot}12 \times 10^4 = 2{\cdot}98 \times 10^4\gamma$ in e.m.u. Therefore we get $9{\cdot}30 \times 10^3 \times 2{\cdot}98 \times 10^4 \times 2{\cdot}63 K_m = 6{\cdot}3$, or

$$K_m = 8{\cdot}6 \times 10^{-9} \text{ sec-cm}^{-1}. \tag{4.3}$$

Previous theoretical work by SCHUSTER (CHAPMAN and BARTELS, 1951, p. 754), (who did not have the detailed tidal theory to work with, and consequently estimated $V_2(a+100) \cong V_2(a)$, i.e. too small by about a factor 200) gave $K_m = 3 \times 10^{-6}$ e.m.u., almost certainly too high. Our value may be compared with an empirical value $K_m \cong 10^{-7}$ e.m.u. for the ionosphere quoted by MITRA (1952, p. 382) as required by radio measurements.

Incidentally, the value (4.3) agrees quite well with the theoretical value 4×10^{-9} e.m.u. (up to 150 km) calculated by BATES and MASSEY (1951) in a painstaking investigation of the conductivity of the ionosphere. Therefore we find no difficulty in explaining the S_q field by means of a dynamo theory with only the reasonable magnification 170. This disagrees with the conclusions of those authors, who found about a 6000-fold magnification necessary to reconcile the strength of the observed S_q field with the dynamo theory. This disagreement is due to the fact that they used the value $(K_m)_{\text{chap}} = 2{\cdot}5 \times 10^{-5}\, l_s^{-1}$ e.m.u. quoted from CHAPMAN as the total conductivity of the ionized layer necessary to give the observed S_q field. Here l_s is the ground-to-ionosphere velocity magnification. On the other hand, we obtained (4.3), or $K_m = 1{\cdot}5 \times 10^{-6}\, l_s^{-1}$ e.m.u. for this quantity. Apart from the general fact that CHAPMAN did not use a detailed tidal theory to derive his current system, we do not know the reason for this disagreement.

5. CONCLUSIONS

Thus a detailed calculation does not remove the paradox presented by equations (2.10) and (2.11). Two possible resolutions of this puzzle will be suggested below. First we want to dismiss the possibility that the explanation is furnished by the thermal convective motions. It seems reasonable that these motions might be equally important as the diurnal tidal motions, but not plausible that they could enhance the purely tidal U_1 a thousandfold.

One possible explanation is to take (4.2) seriously, that is, to look for a resonating diurnal mode in the PEKERIS theory. This would occur if the atmosphere possessed a free oscillation of period near to 24 hr. That theory shows that this will be so if, and only if, the equivalent depth H of the assumed diurnal mode with $\Psi \cong Y_1^1$ nearly coincides with one of the $H^{(j)}$ $(j = 1, 2, \cdots)$ of the free atmosphere. These latter are the equivalent depths which make the associated vertical velocities w_j' of the freely oscillating atmosphere just vanish at the ground. For any given temperature distribution, the $H^{(j)}$ are determined by numerically solving the χ equation (PEKERIS, 1937, equation (12)). For "atmosphere B," PEKERIS finds $H^{(1)} = 10{\cdot}43$ km, $H^{(2)} = 8{\cdot}192$ km. As far as we know, it is unknown whether or not there exists such a diurnal mode with H near to one of these two values. From $V_2(a) \cong 3V_1(a)$ and (4.2) one sees that the ground-to-100 km magnification of this resonating mode would have to be about $7 \times 170 \cong 1200$. Even if there were no resonating diurnal mode, it would be sufficient for $V_1(r)$ to be somehow magnified on going from $r = a$ to $r = a + 100$. But it is expected that this magnification occurs only for a resonating mode.

The second possible explanation seems more promising. As mentioned in Section 3, the rotation of the atmosphere rigidly along with the earth has as a result that the actual semidiurnal and diurnal horizontal velocities rise to high but finite maxima in the polar regions and around latitudes 30°N and S, respectively. Assume that we are going to evaluate the surface integral part of (2.5) numerically (as would probably have to be done anyway for $\mathbf{v}$ the solution of the full non-linear tidal problem). Divide the angular domain into the elements $(d\theta\, d\phi)_j$ $(j = 1, 2, \cdots N)$ (say, square, 1° on a side) and let $\mathbf{I}(\theta_j, \phi_j)$ denote the integrand of the angular integral in (2.5) evaluated at the centre θ_j, ϕ_j of the jth square. Then $\mathbf{B}_i \cong \sum_{j=1}^{N} \mathbf{I}_i(\theta_j, \phi_j) \sin\theta_j (d\theta\, d\phi)_j$ where $i = 1$ and 2 refer to the use of the accurate "non-linear" diurnal and semidiurnal velocities, respectively, in (2.5). Now the $\mathbf{I}_2(\theta_j, \phi_j)$ would be large around the poles but it is just here that the large $\mathbf{I}_2(\theta_j, \phi_j)(d\theta\, d\phi)_j$ would enter the sum with the small weights $\sin\theta_j \cong \sin 0 = 0$. Therefore the high maxima around the poles need not enhance too much U_2 calculated from $\mathbf{v}_{2\text{cutoff}}$ in Section 3. On the other hand, for the diurnal mode, the terms $\mathbf{I}_1(\theta_j, \phi_j)$ $(d\theta\, d\phi)_j$, large around $\theta = 60°$ and 120°, enter the sum provided with the large weights $\sin\theta_j \cong \sin 60° = 0{\cdot}87$. Therefore the U_1 calculated in Section 3 for $\mathbf{v}_{1\text{cutoff}}$ could be greatly enhanced by using instead the realistic diurnal tidal velocities.

A further relevant point can be made here. If the atmosphere does not follow the earth rigidly but lags somewhat behind* as r increases, U_1 might be expected to be even more enhanced relative to U_2. The reason is the following. A distinction should be made in the TAYLOR–PEKERIS theory between ω, the angular frequency of the rotating earth, and ω', the angular frequency of the rotating atmosphere at any point. For simplicity, assume that ω' is a function only of r. Then $\omega'(a) = \omega$ but $\omega'(r) \leq \omega$, $r > a$, allowing for the possibility that the atmosphere lags the earth. The frequency σ is defined in terms of ω, e.g. $\sigma = \omega$ and 2ω for diurnal

* Work to be published elsewhere shows that such an atmospheric lag can be invoked to explain the shift westward of the "cosmic ray equator" from the geomagnetic equator.

VOL. 16 1959

and semidiurnal modes, respectively. But the Coriolis force terms in the motion equations (PEKERIS, 1937), (1) and (2) involve $\omega'(r)$. Consequently the horizontal velocities are provided with the factor $(f^2 - \cos^2\theta)^{-1}$ where $f \equiv \sigma/2\omega'(r)$. Thus the "linearized" horizontal velocities become infinite at $\theta = \cos^{-1}(\pm f)$. If $\omega'(a + 100) < \omega$, this has the effect of increasing f and thus of shifting the north and south latitudes of singularity for the theory with $\omega'(r) \equiv \omega$ northward and southward, respectively. Hence if we assume $\omega'(a + 100) < \omega$, the "linearized" semidiurnal horizontal velocities are no longer infinite anywhere, but the "linearized" diurnal horizontal velocities continue to become infinite, at latitudes north of 30°N and south of 30°S. Thus the maxima of the actual semidiurnal horizontal velocities may be expected to be lower than they are for $\omega'(r) \equiv \omega$; whereas the maxima of the actual diurnal velocities may be expected to remain roughly as high as they are for $\omega'(r) \equiv \omega$, the latitudes of these maxima merely being somewhat shifted.

If a method could be devised of measuring the strength of the various Fourier components $\sigma = \omega, 2\omega, 3\omega$, etc., of the tidal velocities in the upper atmosphere, it would be interesting to check this prediction of tidal theory that a semidiurnal motion prevails in the polar regions, whereas a diurnal motion preponderates around latitudes 30°N and S.

L.
6
59

APPENDIX

The tidal velocities for the diurnal and semidiurnal modes (3.1) are (PEKERIS, 1937, equation (10))

$$\begin{aligned} u_1 &= \frac{i}{4}\frac{V_1(r)}{(\tfrac{1}{4} - \cos^2\theta)}\left[\frac{\partial}{\partial\theta} Y_1^1 + 2\cot\theta\, Y_1^1\right] \\ v_1 &= -\frac{1}{4}\frac{V_1(r)}{(\tfrac{1}{4} - \cos^2\theta)}\left[2\cos\theta\frac{\partial}{\partial\theta} Y_1^1 + \frac{1}{\sin\theta} Y_1^1\right] \\ w_1 &= w_1(r)\, Y_1^1 \end{aligned} \tag{A1}$$

and

$$\begin{aligned} u_2 &= \frac{i}{2}\frac{V_2(r)}{\sin^2\theta}\left[\frac{\partial}{\partial\theta} Y_2^2 + 2\cot\theta\, Y_2^2\right] \\ v_2 &= -\frac{1}{2}\frac{V_2(r)}{\sin^2\theta}\left[\cos\theta\frac{\partial}{\partial\theta} Y_2^2 + \frac{2}{\sin\theta} Y_2^2\right] \\ w_2 &= w_2(r)\, Y_2^2 \end{aligned} \tag{A2}$$

respectively, where $V_i(r) \equiv Q_i(r)/r\omega$. Then the polar coordinate components of

the corresponding dynamo currents $\mathbf{j}_i \equiv \kappa/c(\mathbf{v}_i \times \mathbf{B}_0)$ are, respectively,

$$(j_1)_r = -\frac{\kappa}{c}\frac{M}{r^3}\frac{V_1(r)}{4}\frac{1}{(\frac{1}{4}-\cos^2\theta)}\left[2\cos\theta\sin\theta\frac{\partial}{\partial\theta}Y_1^1 + Y_1^1\right]$$

$$(j_1)_\theta = \frac{\kappa}{c}\frac{M}{r^3}\frac{V_1(r)}{4}\frac{1}{(\frac{1}{4}-\cos^2\theta)}\left[4\cos^2\theta\frac{\partial}{\partial\theta}Y_1^1 + 2\cot\theta\ Y_1^1\right] \tag{A3}$$

$$(j_1)_\phi = i\frac{\kappa}{c}\frac{M}{r^3}\frac{V_1(r)}{4}\left[\frac{1}{(\frac{1}{4}-\cos^2\theta)}\left(2\cos\theta\frac{\partial}{\partial\theta}Y_1^1 + 4\cos\theta\cot\theta\ Y_1^1\right) + 4i\frac{w_1(r)}{V_1(r)}\sin\theta\ Y_1^1\right]$$

and

$$(j_2)_r = -\frac{\kappa}{c}\frac{M}{r^3}V_2(r)\frac{1}{\sin^2\theta}\left[\frac{1}{2}\sin\theta\cos\theta\frac{\partial}{\partial\theta}Y_2^2 + Y_2^2\right]$$

$$(j_2)_\theta = \frac{\kappa}{c}\frac{M}{r^3}V_2(r)\frac{1}{\sin^2\theta}\left[\cos^2\theta\frac{\partial}{\partial\theta}Y_2^2 + 2\cot\theta\ Y_2^2\right] \tag{A4}$$

$$(j_2)_\phi = i\frac{\kappa}{c}\frac{M}{r^3}V_2(r)\left[\frac{1}{\sin^2\theta}\left(\cos\theta\frac{\partial}{\partial\theta}Y_2^2 + 2\cos\theta\cot\theta\ Y_2^2\right) + i\frac{w_2(r)}{V_2(r)}\sin\theta\ Y_2^2\right].$$

VOL. 16 1959

For any vector field $\mathbf{A}$, the circular components A_σ are given directly from the polar components:

$$T\begin{pmatrix}A_r\\A_\theta\\A_\phi\end{pmatrix} = \begin{pmatrix}A_+\\A_-\\A_0\end{pmatrix}$$

by application of the matrix

$$T = \begin{pmatrix}\sin\theta\ e^{i\phi} & \cos\theta\ e^{i\phi} & ie^{i\phi}\\ \sin\theta e^{-i\phi} & \cos\theta e^{-i\phi} & -ie^{-i\phi}\\ \cos\theta & -\sin\theta & 0\end{pmatrix}.$$

Thus the $(j_i)_\sigma$ for the unmodified PEKERIS theory are obtained by applying T to the components (A3) and (A4). The $(j_i)_\sigma$ for the cutoff velocities given by (3.3) with (3.4) and (3.5) are obtained by supressing the factors $(\frac{1}{4}-\cos^2\theta)^{-1}$ and $\sin^{-2}\theta$ in (A3) and (A4), respectively.

To attain the forms given by (3.4) and (3.5), which are linear in spherical harmonics, a certain amount of manipulation is necessary. To this end the following formulae are useful

$$\sin\theta\, e^{i\phi}\, Y_l^{m-1} = \sqrt{\left\{\frac{(l-m)(l-m+1)}{(2l-1)(2l+1)}\right\}}Y_{l-1}^m - \sqrt{\left\{\frac{(l+m)(l+m+1)}{(2l+1)(2l+3)}\right\}}Y_{l+1}^m$$

$$\sin\theta\, e^{-i\phi}\, Y_l^{m+1} = -\sqrt{\left\{\frac{(l+m)(l+m+1)}{(2l-1)(2l+1)}\right\}}Y_{l-1}^m + \sqrt{\left\{\frac{(l-m)(l-m+1)}{(2l+1)(2l+3)}\right\}}Y_{l+1}^m$$

$$\cos\theta\, Y_l^m = \sqrt{\left\{\frac{(l-m)(l+m)}{(2l-1)(2l+1)}\right\}} Y_{l-1}^m + \sqrt{\left\{\frac{(l-m+1)(l+m+1)}{(2l+1)(2l+3)}\right\}} Y_{l+1}^m,$$

$$\frac{\partial}{\partial\theta} Y_2^2 = 2\cot\theta\, Y_2^2 = -e^{i\phi} Y_2^1$$

$$\frac{\partial}{\partial\theta} Y_1^1 = \cot\theta\, Y_1^1 = -\frac{\sqrt{2}}{2} e^{i\phi} Y_1^0$$

and

$$Y_2^0 Y_2^2 = -\frac{1}{7}\sqrt{\frac{5}{\pi}}\, Y_2^2 + \frac{1}{14}\sqrt{\frac{15}{\pi}}\, Y_4^2$$

$$Y_2^1 Y_2^2 = \frac{1}{14}\sqrt{\frac{5}{\pi}}\, Y_2^1 + \frac{3}{7}\sqrt{\frac{5}{6\pi}}\, Y_4^1$$

$$Y_2^0 Y_2^0 = \frac{1}{2}\sqrt{\frac{1}{\pi}}\, Y_0^0 + \frac{1}{7}\sqrt{\frac{5}{\pi}}\, Y_2^0 + \frac{3}{7}\sqrt{\frac{1}{\pi}}\, Y_4^0.$$

References

Bates D. and Massey H.	1951	*J. Atmosph. Terr. Phys.* **2,** 1.
Chapman S. and Bartels J.	1951	*Geomagnetism* Clarendon Press, Oxford.
Condon E. and Shortley G.	1953	*The Theory of Atomic Spectra* Cambridge University Press.
Mitra S. K.	1952	*The Upper Atmosphere* (2nd Ed.) The Asiatic Society, London.
Pekeris C.	1937	*Proc. Roy. Soc.* A **158,** 650.
Rose M. E.	1955	*Multipole Fields* Chapman and Hall, London.
Spitzer L.	1956	*Physics of Fully Ionized Gases.* Interscience, New York.
Taylor G.	1936	*Proc. Roy. Soc.* A **156,** 378.
Wilkes M.	1949	*Oscillations of the Earth's Atmosphere.* Cambridge University Press.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 274 to 282. Pergamon Press Ltd. Printed in Northern Ireland

The relation between H- and Z-variations near the equatorial electrojet

C. A. Onwumechilli
Physics Department, University College, Ibadan, Nigeria

(*Received* 12 *January* 1959)

Abstract—The parallelism in the variations of the geomagnetic horizontal component (H) and vertical component (Z), so obvious in the mangetograms of Ibadan, has been investigated. During day-time when the effect of the electrojet is dominant, H and Z vary proportionately. Numerous indentations of varying period and amplitude are a prominent feature of day-time variation curves of H and Z at Ibadan. Six hundred of these indentations have been measured and analysed. The indentations attain their maximal values simultaneously in H and Z. Within the limits of experimental error, in 86 per cent of the cases during the day and 54 per cent of those during the night, the indentations also begin and end simultaneously in H and Z.

The ratio r of the amplitude of an indentation in H to the amplitude of the same indentation in Z exhibits a remarkable diurnal variation, being constant for most of the day-time. The average value is $2{\cdot}114 \pm 0{\cdot}021$ during daytime and $4{\cdot}112 \pm 0{\cdot}095$ during the night. The ratio r is surprisingly independent of magnetic disturbance but decreases with increase in the period of duration T. For indentations of short periods T (e.g. not exceeding 15 min) the ratio r equals the ratio R of the diurnal range of H to the range of Z on undisturbed days. It is suggested that the indentations are caused by fluctuations in the quantity and quality of ionizing agents from the sun and that the decrease of r with increase in T is associated with vertical movements of the E-layer of the ionosphere (or the S_q-layer).

VOL.
16
1959

1. Introduction

Ibadan is only $2\frac{1}{2}°$ south of the magnetic equator and is estimated to be 240 km south of the centre of the equatorial electrojet. At this distance the effect of the electrojet on both the variations of H and Z is great. It has been estimated (Onwumechilli, 1959) that the electrojet accounts for half the amplitude of the variation of H, and about four fifths or more of the variation of Z at Ibadan. It would therefore be expected that these variations arising largely from the same cause will have much in common. Investigations have been carried out into the relationships between the diurnal as well as the short-time variations in H and Z.

2. Parallelism of H and Z Variation Curves

Ibadan magnetograms (Fig. 1) give the impression that variations in Z follow those of H even in minute details. The indentations in H-variation curve have corresponding indentations in the Z-curve. On examination of the indentations it is found that H and Z both increase and decrease together.

For a more detailed study we first consider the diurnal variations under quiet conditions. Fig. 2 shows the curves of the average diurnal variation of H and Z on international quiet days for the three seasons—December solstice (D), equinoxes (E) and June solstice (J) during the period of November 1955 to June 1957. The general features of the variations are the same, and H and Z attain their maxima about the same time but the Z-variation reaches deeper minima than H at sunrise and sunset.

The total magnetic field of the earth is a vector in space the end of which traces a curve as the field goes through its diurnal variation. Fig. 3 is the projection of this curve on the HZ-plane (the vertical plane through the magnetic meridian).

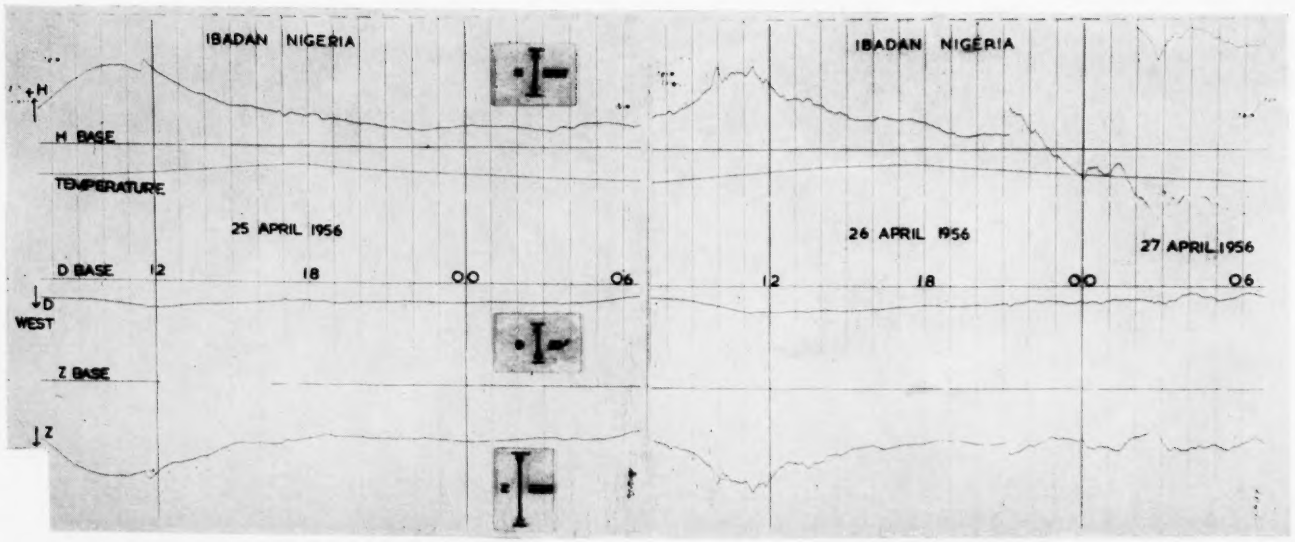


Fig. 1. Sample magnetic record from Ibadan, Nigeria.

VOL.
16
1959

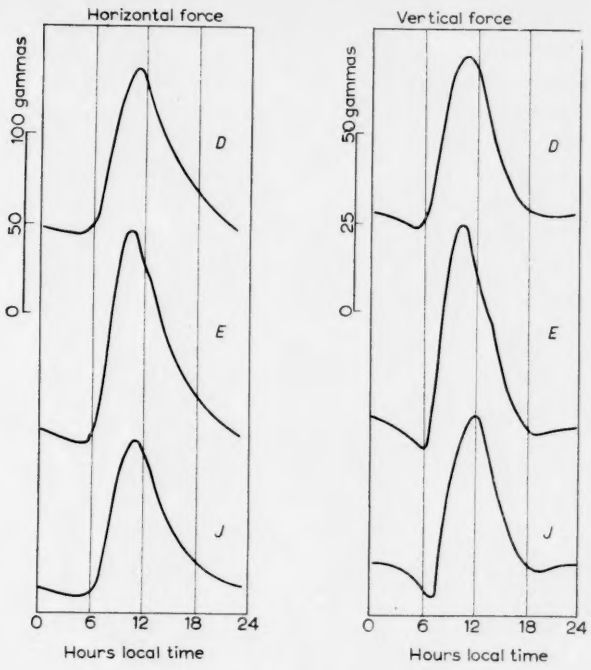


Fig. 2. Showing seasonal means of the solar daily variation of H and Z on international quiet days at Ibadan.

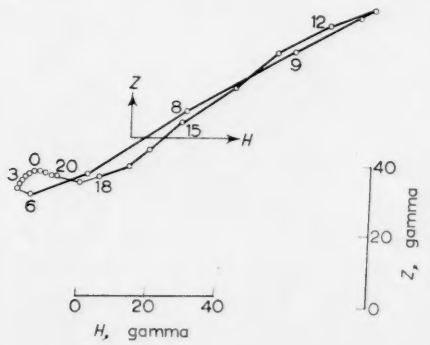


Fig. 3. Showing vector diagram for the solar daily variation of H and Z at Ibadan on international quiet days.

This vector diagram giving the combined variation of H and Z has been drawn for the international quiet days in the 20 months of November 1955 to June 1957. Its needle shape shows that the changes in H are proportional to the changes in Z. The proportional relation, however, does not strictly hold in the morning and in the evening. Both H and Z change very little in the night (2000 to 0500). The vector diagram thus brings out well the parallelism between the variations.

3. Indentations in H and Z Variation Curves

Having observed that indentations in H-variation curves are reproduced in Z-variation, and that they indicate contemporaneous increases (or decreases) in H and Z, investigations have been directed towards ascertaining whether, (a) the fluctuations in H and Z grow to their maxima in the same time, (b) the fluctuations last for the same length of time and (c) whether the fluctuations in H and Z are proportional as in the case of the diurnal variation on quiet days.

For this investigation 600 indentations in H and Z-variation curves at Ibadan have been measured. The H- and Z-scale values of the records are 3·74 and 2·35 gammas/mm respectively so that measuring with a transparent millimetre scale,

Table 1. Showing differences in periods of H- and Z-fluctuations

Number of indentations	Day-time 0730 to 1430		Night-time 1930 to 0330	
	Time of growth t_1	Period of duration T	Time of growth t_1	Period of duration T
In all	309	309	82	82
With no difference	196	201	38	41
With difference $\leq$ 2 min	271	266	48	44

VOL. 16 1959

the accuracy is better than 1·9 and 1·2 gammas in H and Z, respectively. An indentation is not measured unless it has an amplitude of at least 1 mm in both H and Z. An indentation is superposed on the diurnal variation. Thus if h_1 is the ordinate of the curve (from a fixed base line) at the point where it departs from the trend of the diurnal variation and h_2 the ordinate at the point where it returns to the general trend, $\frac{1}{2}(h_1 + h_2)$ is taken as the mean level of the diurnal variation during the period of the indentation. If, therefore, the maximal ordinate (maximum or minimum) within the indentation is h_3, the amplitude of the indentation m is given by

$$m = h_3 - \tfrac{1}{2}(h_1 + h_2). \tag{1}$$

The time t_1 of the growth of any fluctuation is the interval between the ordinates h_1 and h_3; and the time of decay t_2 is the interval between the ordinates h_2 and h_3. The total period of duration T is the interval between the ordinates h_1 and h_2 and equals $(t_1 + t_2)$. The time scale of the records is 4 min/mm, so the periods determined with the millimetre scale should be accurate to better than 2 min. An indentation is not considered for measurement unless its period of duration T is up to 8 min. About a third of the 600 indentations have a duration period of not more than 20 min but some long periods of about 120 min occur.

4. Simultaneity of Fluctuations in H and Z

In most cases the indentations begin, reach their maximal values, and end at the same time in H and Z. The times of growth and periods of duration of 309 indentations in H and Z between the hours of 0730 and 1430 GMT and 82 indentations between the hours of 1930 and 0330 GMT have been compared in Table 1

(Greenwich Mean Time is only 15 min behind the local time of the observatory). The times of growth of the same indentation in H and Z are equal in 63·4 per cent of the cases during the day and 46·3 per cent of the cases during the night. In 87·7 per cent of the day-time cases and 58·5 per cent of the night-time cases the difference is not more than 2 min—the possible error in the measurements. The period of duration is the same in 65 per cent of the cases during the day and 50 per cent of the cases during the night, and the difference is not more than 2 min in 86 per cent and 53·6 per cent of the cases during the day and night, respectively.

It may therefore be concluded that within the errors of the measurments, the indentations in H and Z during the day take the same time to grow and last for the same period. Since it is further observed that in day-time nearly all the maxima occur at the same time, it means that the fluctuations also start and end at the same time. During the night, on the other hand, the chance that a fluctuation grows or lasts for the same time in H and Z is the same as the chance that the times differ.

5. Amplitudes of Indentations in H and Z Diurnal Variation Curves

The amplitudes m of indentations determined according to equation (1), give the ratio r of the amplitude of an indentation in H to the amplitude of the same indentation in Z. Each of the 600 ratios is grouped into its hour of occurrence, being the hour nearest the time of its maximum; thus the group for 0800 includes r for all indentations reaching their maxima between 0730 and 0830. The hourly mean ratios are plotted in Fig. 4. The dotted vertical lines indicate the limits of the errors expected in the determinations of the means. For the hours of 0700 to 1400 the errors are too small to be indicated in the same way. The single indication below 1100 hours is the mean for the hours and could stand for any of them.

An outstanding feature of Fig. 4 is that the ratios of the amplitudes of fluctuations is much greater in the night than in the day. The mean for 12 hr of daylight, 0430 to 1630, is $2{\cdot}114 \pm 0{\cdot}021$, and for the 12 hr of darkness, 1630 to 0430, is $4{\cdot}112 \pm 0{\cdot}095$. For the hours of 0800 to 1400 the ratio r is practically constant (mean $2{\cdot}145 \pm 0{\cdot}021$). At sunrise (0500 and 0600) r is smallest and at sunset (1500 and 1600) r is slightly larger than the day-time mean.

The 342 values of r from 0800 to 1400 hours have been used for a more detailed study of the ratios. Fig. 5 giving the frequency distribution of the 342 values suggests that the distribution is fairly normal.

6. Dependence of the Ratio r on the Period of Duration T

The 342 indentations for the hours of 0800 to 1400 have been arranged into seven groups A, B, C, D, E, F and G in increasing periods of duration (Table 2). It is found that r decreases as the period of duration T increases (Fig. 6). The difference between the mean ratios of groups A and C is just over three times the expected error, and as the periods increase in groups D–G the margin is steadily widened. The first and last groups differ by ten times the expected error in the mean. The usual statistical criterion for significance—that the difference must be at least three times the error—is therefore satisfied.

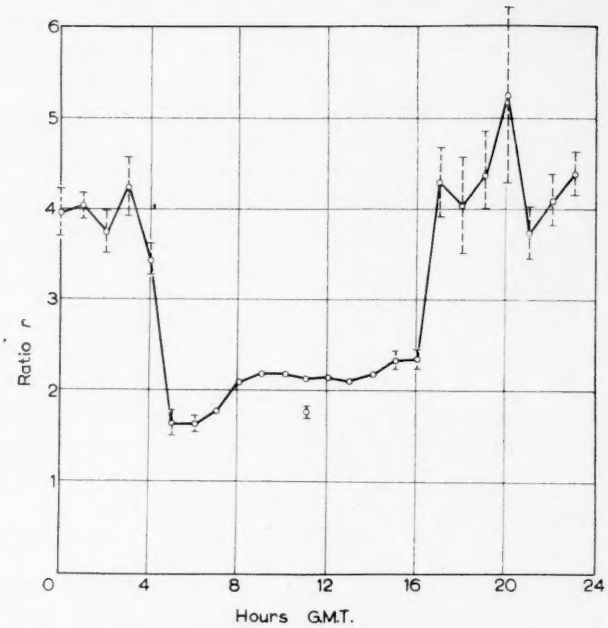


Fig. 4. Showing the diurnal variation of the ratio r of the amplitude of a fluctuation in H to the amplitude of the same fluctuation in Z.

VOL. 16 1959

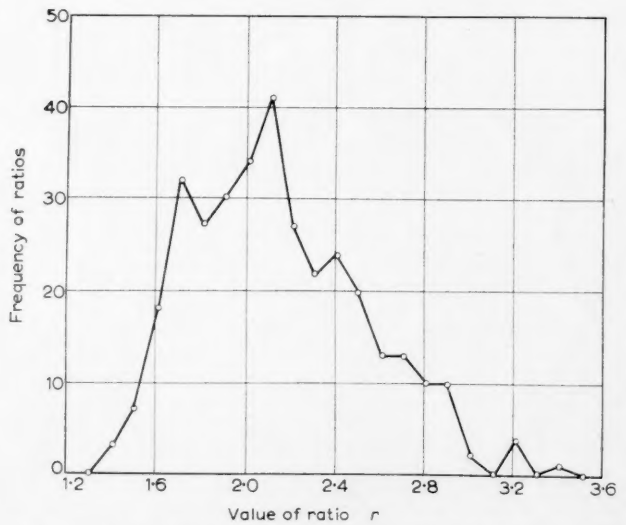


Fig. 5. Showing the frequency distribution of the ratio r for 0800 to 1400 hours GMT.

7. Dependence of the Ratio r on Disturbance

In the measurements, no discrimination is made between quiet and disturbed conditions. It is therefore important to know how r varies with disturbance. The indentations are not numerous at night and may be entirely absent on quiet nights.

Table 2. Showing the variation of the ratio r with period of duration T

Group	A	B	C	D	E	F	G
Period T min	<11	11–16	17–20	21–30	31–40	41–60	>60
Count N	62	87	41	46	47	32	27
Mean $\bar{r}$	2·381	2·295	2·210	2·083	1·911	1·890	1·833
Error of mean	0·049	0·043	0·040	0·055	0·044	0·055	0·048

Hence it is also necessary to find out if the higher ratios at night are caused by the higher level of disturbance necessary to produce indentations at night.

For this investigation the 342 values of r for 0800 to 1400 hours are divided into ten groups according to the 3 hr planetary K-indices under which the indentations

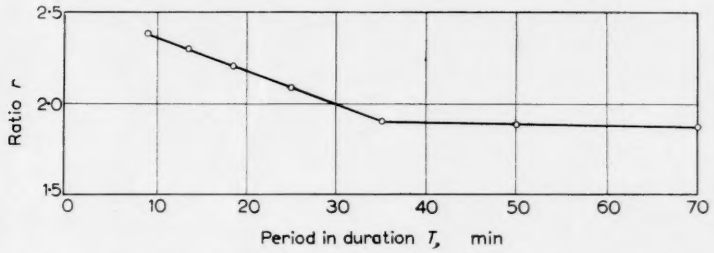


Fig. 6. Showing the variation of the ratio r with the period of duration T.

occurred. The means of the groups plotted in Fig. 7 show that the variation of r with K_p has no easily recognizable pattern.

To prove that this surprising result is not caused by uneven distribution of periods of duration among the K_p index groups, the 87 values of r of period 11 to

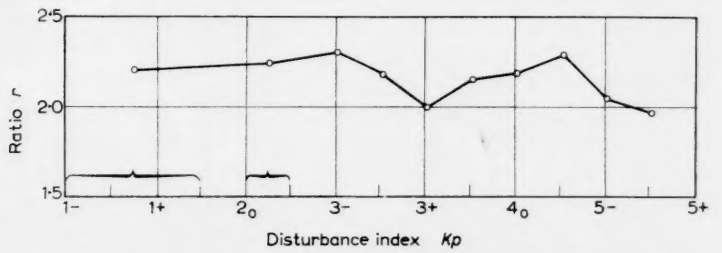


Fig. 7. Showing the variation of the ratio r with the K_p-disturbance index.

16 min (group B of Table 2) are subdivided into three, according to their K_p indices. The result in Table 3 again suggests that r does not depend on disturbance.

8. Comparison of r with the Ratio of Daily Ranges R

In the 2 months under study (April and May 1958) 36 undisturbed days, mostly with magnetic character figure $C_p < 1$ and with average $C_p = 0{\cdot}67$, are

selected. The absolute maximum and two minima (one in the morning and the other in the evening) are measured in each record for both H and Z. The highest mean in an hour interval (not necessarily coinciding with hour marks), the smallest mean in an hour interval in the morning and the smallest in the evening are also

Table 3. Showing the dependence of r on disturbance

K_p group	Count	Mean r	Error of mean
0_0 to 3_0	27	2·437	0·064
3_+ to 4_-	24	2·063	0·063
4_0 to 5_-	31	2·386	0·081

determined for H and Z in each day. From these, several slightly differing daily ranges may be computed. The absolute range M_1 is the absolute maximum minus the absolute minimum. The mean daily range M_2 is the difference between the maximum and minimum hourly means. The adjusted absolute range M_3 is the absolute maximum minus the mean of the morning and evening minima. The

Table 4. Showing the average of the ratio of the daily range of H to the range of Z on 36 undisturbed days (mean $C_p = 0·67$)

	R_1	R_2	R_3	R_4
Mean ratio	2·298	2·374	2·361	2·433
Error of mean	0·084	0·088	0·075	0·085

adjusted mean daily range M_4 is the maximum hourly mean less the average of the morning and evening minimum hourly means. On dividing the range of H by the corresponding range of Z, four ratios R_1, R_2, R_3 and R_4 are obtained corresponding to the ranges defined above. The average of the thirty-six ratios for each type of range is given in Table 4.

It is seen that no matter how the range is defined, the ratio R, of H to Z daily range, equals the ratio r of the amplitudes of fluctuations in H to the amplitudes of corresponding fluctuations in Z during day-time. This is particularly true of fluctuations of comparatively short period of duration, say, 8–16 min (Table 2).

9. Discussions of Results

One striking thing about the investigations is the significant difference between night and day in respect of each phenomenon considered. Perhaps because the electrojet is absent at night, H and Z vary little, and changes in H are no longer proportional to changes in Z as is the case in day-time (Fig. 3). This helps to explain the difference between day and night results in the comparisons of periods of H- and Z-fluctuations. The observed indentation is a combination of the actual fluctuation and the diurnal variation. Since the diurnal variations of H and Z

VOL. 16 1959

proceed at different rates at night, their combinations with the fluctuation should reach their maxima at different times. This difference ought to be small since H and Z change very little at night. Therefore it seems possible that other forces are also at work to make periods of H- and Z-fluctuations differ by more than 2 min in 46 per cent of the cases at night as compared with only 14 per cent during the day. It is possible that the predominant cause of and/or medium of production of indentations during the day and night are fundamentally different. For one thing the electrojet effect is absent at night and since it accounts for a larger part of Z- than H-variation at Ibadan the ratio r should be larger (as is found) during the night ($4{\cdot}112 \pm 0{\cdot}095$) than during the day ($2{\cdot}114 \pm 0{\cdot}021$).

The question of the cause of the fluctuations can only be broached at the moment. The merits of some possible causes are examined hereunder in the light of existing knowledge. If geomagnetic variations are produced in the ionosphere by dynamo action, the important variables to focus attention on are the conductivity and the dynamo e.m.f. The dynamo e.m.f. is the vector product of the velocity of charged particles and the absolute value of the geomagnetic field. Nevertheless if the fluctuations are caused by changes in the dynamo e.m.f. we must consider fluctuations in the wind, since the percentage change in the magnetic field is very small.

Changes in the wind direction must be dismissed from the possible cause of indentations because such a change would neither keep r constant for any length of time nor make r equal to R. On the other hand, changes in wind speed can cause proportional fluctuations in H and Z as observed, but only if the wind direction remains constant. It is rather unlikely that throughout the day-time the wind speed would change hundreds of times without being accompanied with changes in direction. Even then changes in wind speed can neither account for the dependence of r on T (Fig. 6), nor the equality of r and R.

We must therefore look into ionospheric conductivity for the cause of the indentations. The important factors in conductivity are the density of ionization and the height of the conducting layer. Since the ratio r of the fluctuation amplitudes equals the ratio R of daily ranges, we expect the changes in conductivity to explain the fluctuations and the diurnal variations in the same way. One basic requirement, therefore, is that while the fluctuation in current strength is in progress, the wind direction should remain reasonably the same as during the major part of the diurnal variation. If the density of ionization fluctuates, the most likely causes are changes in the quality and/or quantity of ionizing radiations from the sun, or in more local cases by the drift of clouds of ionization. These fluctuations are found to be contemporaneous at Ibadan and Katsina—a distance of over 700 km (Onwumechilli, 1959). It is therefore doubtful that localized clouds of ionization can be the answer, even if they are centred on the electrojet. A more likely cause is the fluctuation in the quantity and quality of ionizing radiations from the sun. A burst of ionizing radiations boosts ionospheric conductivity and causes a temporary augmentation of the S_q-current as is now thought to be the case during a solar flare effect (McNish, 1937, Nagata, 1952; Pratap, 1957). It would be easy to detect small fluctuations in ionizing radiations because the conductivity near the electrojet is so much magnified, but in order to

explain the dependence of r on T this suggestion has to be tentatively extended. The observed inverse relation between r and T can be explained by vertical movement of ionization; and in this line of argument it will be required that the radiation becomes harder as S_q is augmented so that the changes in the height of maximum ionization density in the S_q-layer will produce the effect of vertical movement of ionization.

It has been shown (Hirono 1950a, b; Baker and Martyn, 1953) that the great enhancement of ionospheric conductivity within the equatorial electrojet depends on the ultimate inhibition of Hall current by the polarization field set up by the initial flow of the current. It can scarcely be imagined that the conditions for a complete inhibition of the current will remain steadily satisfied throughout the day. It is not unreasonable, therefore, to suppose that the condition for complete inhibition of Hall current fails for short intervals and causes the observed fluctuations in S_q-current. Such may easily explain the dependence of r on T but must not be very much localized and in any case is likely to lack the suddenness with which the fluctuations sometimes occur.

10. Conclusion

The parallelism of H- and Z-variations at Ibadan arises from the fact that both variations are largely controlled by the equatorial electrojet. The parallelism between the short time fluctuations in H and Z in a similar way suggests that they manifest changes in the strength of the electrojet. Whether these changes arise from the internal constitution of the equatorial electrojet or from a wider cause (like the ionizing radiations) to which the electrojet is sensitive should be elucidated by further investigations planned, but the present work has shown that the fluctuations are not likely to arise from the wind. The facts that (a) the ratio of the amplitudes of fluctuations in H and Z equals the ratio of H- and Z-ranges and (b) that the ratio of fluctuation amplitudes is independent of magnetic disturbance imply that during the fluctuations the wind direction in the S_q-layer remains largely the same as during the major part of the diurnal variation, and that this condition continues to hold during disturbances (at least for moderate ones $K_p < 6$).

Acknowledgements—I wish to express my gratitude to my colleagues of Physics Department for their interest and suggestions in the course of this work.

References

Baker W. G. and Martyn D. F.	1953	*Phil. Trans.* A **246,** 281.
Hirono M.	1950a	*J. Geomagn. Geoelect., Kyoto* **2,** 1.
Hirono M.	1950b	*J. Geomagn. Geoelect., Kyoto* **2,** 113.
McNish A. G.	1937	*Terr. Magn. Atmos. Elect.* **42,** 109.
Nagata T.	1952	*J. Geophys. Res.* **57,** 1.
Onwumechilli C. A.	1959	*J. Atmosph. Terr. Phys.* **13,** 222, 235.
Pratap R.	1957	*J. Geophys. Res.* **62,** 581.

VOL. 16 1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 283 to 290. Pergamon Press Ltd. Printed in Northern Ireland

The absorption of short radio waves in the *D*-, *E*- and *F*-regions of the ionosphere

J. D. WHITEHEAD
University of Otago, Dunedin, New Zealand

(*Received* 20 *January* 1959)

Abstract—The published values of the noon absorption of radio waves of frequencies 2, 2·4, 4 and 4·8 Mc/s measured at Slough from 1947 to 1953 have been analysed. From the absorptions of 2 and 2·4 Mc/s waves, it has been deduced that (a) the electronic collision frequency ν_0 at the height of the maximum electron density in the *E*-region when the sun's rays are incident vertically on the ionosphere is $(1{\cdot}9 \pm 0{\cdot}1) \times 10^4$ per sec; and (b) the normal absorption in the *D*-region increases with increasing sunspot number, whereas the additional absorption which occurs on certain winter days and arises in the *D*-region may decrease with increasing sunspot number.

The 4 and 4·8 Mc/s waves were reflected from the *F*-region at noon. The absorption in the *F*-region is calculated from the known total absorption and the calculated absorption in the *D*- and *E*-regions. The collision frequency in the *F*-region may then be found. At the height of reflection of 4 Mc/s waves (150 to 180 km) it is $(3{\cdot}6 \pm 0{\cdot}6) \times 10^3$ per sec, and at the height of reflection of 4·8 Mc/s waves (180 to 200 km) it is $(3{\cdot}0 \pm 0{\cdot}6) \times 10^3$ per sec. The significance of these results is discussed.

1. Introduction

When radio waves are propagated through the ionosphere they are absorbed because the free electrons make collisions and loose the ordered vibrational energy given them by the waves. The absorption of radio waves reflected at vertical incidence may be measured in terms of the reflection coefficient ρ of the ionosphere. The absorption is given by

$$-\ln \rho = \int \kappa \, ds$$

the integral being taken over the path s of the wave. κ is the absorption coefficient and depends on the electronic collision frequency, the radio frequency, the electron density and the magnitude and direction of the earth's magnetic field.

The absorption occurs in various parts of the ionosphere. The purpose of this paper is to show how, from the noon values of $-\ln \rho$ measured at Slough, for the ordinary ray on frequencies 2, 2·4, 4 and 4·8 Mc/s, it is possible to separate the contributions to the absorption into the *D*-, *E*- and *F*-region absorptions. Values of the electronic collision frequencies in the *E*- and *F*-regions may then be found.

2. The Absorbing Regions

Let us examine those parts of the ionosphere where the absorption may occur. For radio waves reflected from the *E*-region appreciable absorption may take place in three regions:

(a) Near the level of reflection. Whitehead (1956) showed that a major part of this absorption occurs within a few hundred metres of the level of reflection.

(b) In the *D*-region where the refractive index is nearly unity and the electronic collision frequency is much smaller than the angular frequency of the radio waves.

(c) In the region about the level where the electronic collision frequency ν equals the angular frequency p, of the radio waves.

Below this last region, the collision frequency becomes so large that the few free electrons present do not have time between collisions to extract appreciable energy from the radio waves. Therefore the absorption is very small. For radio waves reflected from the F-region, absorption may occur near the reflection level, in the E- and D-regions and in the "$\nu = p$" region. In each of the absorbing regions, the actual absorption is a function of the frequency of the incident radio waves.

Under the conditions existing in the D-region, BOOKER (1935) showed that the absorption of the ordinary ray is proportional to $(f + f_L)^{-2}$, where f is the radio-frequency and f_L is the longitudinal component of the gyro-frequency.

The exact way in which the absorption varies with frequency in the "$\nu = p$" region, depends on the electron density distribution around this level. If the electron density were constant, the absorption in the region would be proportional to $(f + f_L)^{-1}$ (BOOKER, 1935). On the other hand, if the electron density were increasing much more rapidly with height than the collision frequency was decreasing, the total absorption would be proportional to $(f + f_L)^{-2}$.

Now it is known that on certain winter days there are formed abnormally low layers of ionization capable of partial and total reflection of radio waves (GNANALINGHAM and WEEKES, 1954; DIEMINGER, 1952; GREGORY, 1956), and it is for these days that we must look at the experimental results to see whether there is evidence for appreciable absorption proportional to $(f + f_L)^{-1}$. In fact, WHITEHEAD (1957) showed that the additional absorption, which occurs on days with the abnormally low layers present, is proportional to $(f + f_L)^{-2}$. Therefore we may conclude that at other times also, when the low layers are absent, the absorption below the E-region is proportional to $(f + f_L)^{-2}$: this absorption will be referred to as the D-region absorption.

The absorption in the E-region of radio waves reflected from inside the region or traversing the region may be calculated by assuming that the E-region is a CHAPMAN (1931) region, for it is known (MITRA, 1952) that in the E-region critical frequencies and virtual heights behave almost as though the E-region were a Chapman region. The absorption is then proportional to $H\nu_0$, where H is the scale height and ν_0 the electronic collision frequency at the height of maximum electron density when the sun's rays are incident vertically on the ionosphere.

The absorption in the F-region of waves reflected by the region is proportional to the collision frequency and the reciprocal of the electron density gradient, both taken at the level of reflection. We do not know how the collision frequency varies with height and it is not therefore possible to calculate how the absorption varies with frequency. Once the absorption is known, however, and the electron density gradient found from the N:h profile, it is possible to calculate the collision frequency at the reflection level in the F-region. The results are discussed in Section 5.

3. THE COLLISION FREQUENCY IN THE E-REGION

During the period 1947–1953, 2 and 2·4 Mc/s waves were reflected from the E-region at noon. However, because the absorption varies very rapidly when the

VOL. 16 1959

radio-frequency is close to the critical frequency, results taken when the E-region critical frequency was below 2·7 Mc/s were not used in the following analysis.

On the assumption that the E-region is a Chapman region, the E-region absorption of 2 Mc/s waves is given approximately by*

$$1{\cdot}32\,\frac{H\nu_0}{10^5}\,(1 - 0{\cdot}0013R)\,\cos^{0{\cdot}833}\chi$$

and that of 2·4 Mc/s waves by

$$1{\cdot}50\,\frac{H\nu_0}{10^5}\,(1 - 0{\cdot}0013R)\,\cos^{0{\cdot}833}\chi$$

where χ is the sun's zenith angle and R is the sunspot number.

The ratio of these two E-region absorptions is 0·88 : 1. As the ratio of the D-region absorptions for these two frequencies is known to be 1·27 : 1 (taking

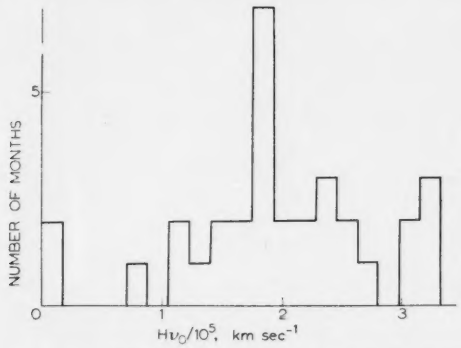


Fig. 1. Histogram of the values of $H\nu_0$ calculated from the monthly mean absorptions of 2 and 2·4 Mc/s waves. H is the scale height of the E-region and ν_0 the electronic collision frequency at the maximum electron density of the E-region when the sun's rays are incident vertically on the ionosphere.

$f_L = 1{\cdot}2$ Mc/s), it is possible, from the observed values of $-\ln\rho$ for the two frequencies, to calculate the E-region absorptions and therefore a value of $H\nu_0$. During winter, the E-region absorption is small and this leads to a large spread in the values for $H\nu_0$. The winter results may be confused also, because in winter reflection may take place from sharp ledges in the E-region, rather than from the Chapman region (Whitehead, 1959). Therefore only the results from April to August were used to calculate the average value of $H\nu_0$. The histogram of values of $H\nu_0$ is plotted in Fig. 1. The mean is $(1{\cdot}9 \pm 0{\cdot}1) \times 10^5$ km/sec. Taking H as 10 km, this gives a mean value of $\nu_0 = (1{\cdot}9 \pm 0{\cdot}1) \times 10^4$ per sec.

It is of interest to see whether the values of $H\nu_0$ vary either with sunspot number R or the cosine of the sun's zenith angle χ. A variation with sunspot number would indicate that ν_0, and with it the height of the E-region, varies through the solar cycle; a variation with $\cos\chi$ would indicate a departure from the Chapman region behaviour.

* These equations were found by numerical integration of the absorption coefficient calculated to include the effect of the earth's magnetic field, and then arranging the parameters to provide the best fit over the required ranges of $\cos\chi$ and R.

In Fig. 2, Hv_0 is plotted against the sunspot number. Each point represents the median of 5 months' results. The variation is certainly less than 20 per cent indicating that the height of the E-region changed by less than 2 km through the sunspot cycle. This method is far more sensitive than direct observations would be.

In Fig. 3, Hv_0 is plotted against cos χ. The winter results are here included. Each point represents the median of 5 months' results. The constancy of Hv_0 justifies the assumption that the E-region behaved very much like a Chapman region.

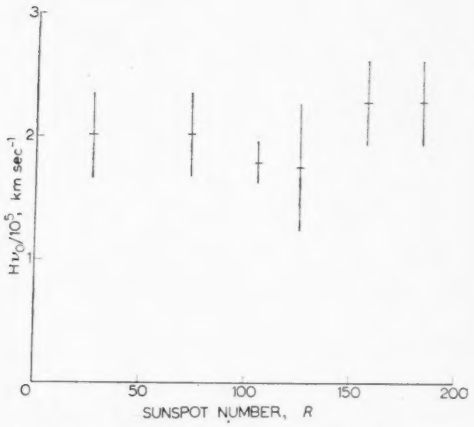


Fig. 2. Plot of Hv_0 against the sunspot number R. The constancy of Hv_0 shows that the height of the E-region does not vary with R.

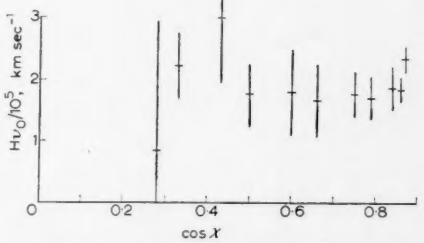


Fig. 3. Plot of Hv_0 against the cosine of the sun's zenith angle χ. The constancy of Hv_0 shows that the E-region behaves almost like a Chapman region.

4. The D-Region Absorption

The D-region absorption may be calculated using either the 2 Mc/s or 2·4 Mc/s total absorption, and the value of Hv_0 is used to calculate the E-region absorption. Except when the E-region critical frequency fell below 2·7 Mc/s both frequencies were used to calculate the D-region absorption. This quantity may be expressed in terms of the parameter D, the D-region absorption at a frequency f being $D(f + f_L)^{-2}$. In Fig. 4, values of D are plotted for each month from 1947–1953. Three points become obvious:

(a) The value of D decreased through the period along with the sunspot number which was a maximum in 1947. Excluding winter results, the dependence of D upon the sunspot number is given by $D = 1{\cdot}4\,(1 + 0{\cdot}011R) \times 10^{13}$ nepers/sec^2.

(b) During the winters of 1948–1949, 1950–1951, 1951–1952, 1952–1953, D was much greater than during the rest of the year.

(c) The peak values of D reached during the winter do not seem to depend on the sunspot number.

It seems that the normally present D-region ionization is attributable to the sun. However, as D shows no variation with cos χ (if we exclude the winter months of course), it seems unlikely that the ionization is entirely due to photon radiation

VOL. 16 1959

from the sun. The sharp gradients of ionization which do exist in the D-region (GREGORY, 1956; SEDDON, 1958) suggest some type of particle ionization.

The additional D-region absorption in excess of the normal component appears to decrease with increasing sunspot number. It will be of great interest to see how the low layers themselves behave through a solar cycle.

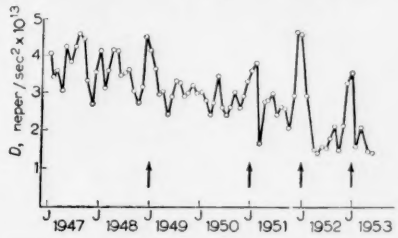


Fig. 4. The monthly mean noon values of D plotted for the period 1947–1953. The D-region absorption of waves of frequency f is $D(f+f_L)^{-2}$.

5. THE F-REGION ABSORPTION

It is now possible, knowing the D-region absorption and Hv_0 for the E-region, to calculate the absorption in the F-region of waves reflected by it. The published results for $-\ln \rho$ on 4 Mc/s and 4·8 Mc/s were used.

These results must first be corrected. In a recent paper, the writer (WHITEHEAD, 1958) has shown that the amplitude of radio waves propagated from a point source vertically into the ionosphere falls off inversely as $P' \sin^2 \theta + P \cos^2 \theta$, where P' and P are the group and actual paths, respectively, and θ is the angle the earth's magnetic field makes with the vertical. The Slough results were calculated by assuming that the amplitude falls off inversely as P' (PIGGOTT, 1953). WHITEHEAD showed that the correct value for the absorption is given by

$$-\ln \rho = -\ln \rho_m - \ln\left\{1 - \left(\frac{P' - P}{P'}\right)\cos^2\theta\right\} + \ln\left\{1 - \left[\frac{P' - P}{P'}\right]_{\text{cal}}\cos^2\theta\right\}$$

where the values of P' and P in the square brackets refer to the group and actual paths during the absolute calibration, and ρ_m is the apparent reflection coefficient.

The third term on the right-hand side of this equation would be about 0·1 neper. The second term must be calculated using the $N : h$ profiles given by THOMAS *et al.* (1957). The total correction to $-\ln \rho$ for these F-region reflections varied from zero (in winter) to 0·27 neper in summer at sunspot minimum. The correction for the E-region reflections is less than 0·05 neper and is neglected.

The absorption in the F-region was now calculated. If the collision frequency of electrons in the F-region does not vary with height, the absorption in F-region will be proportional to dh/dN, the reciprocal of the electron density gradient.

Values of dh/dN for the heights of reflection of 4 Mc/s and 4·8 Mc/s waves were found from the N:h profiles for the ionosphere above Slough (THOMAS *et al.* (1957). Unfortunately, the N:h profiles have only been calculated for 6 months, January ($R = 101{\cdot}6$), September ($R = 51{\cdot}3$) and July ($R = 91{\cdot}0$) 1950; and December ($R = 2{\cdot}5$), September ($R = 19{\cdot}3$) and June ($R = 21{\cdot}8$) 1953. Mean noon values of dh/dN were found for each month, international disturbed days being omitted.

Simple empirical formulae in $\cos\chi$ and R were derived to fit the results. These give, at the level of reflection of the 4 Mc/s waves,

$$\frac{dh}{dN} = 75(1 - 0{\cdot}005R)\cos\chi \text{ cm}^4\text{/electron}$$

and at the level of reflection of 4·8 Mc/s waves

$$\frac{dh}{dN} = 60(\cos\chi - 0{\cdot}12) \text{ cm}^4\text{/electron.}$$

Extrapolation was regarded as being reliable up to $R = 150$. The variation with $\cos\chi$ is a purely seasonal one and the formulae given do not fit diurnal

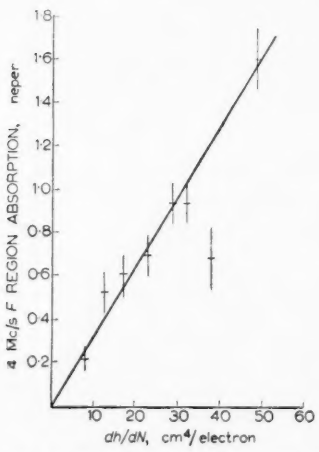


Fig. 5. The F-region absorption of 4 Mc/s waves plotted against dh/dN, the reciprocal of the electron density gradient at the reflection level.

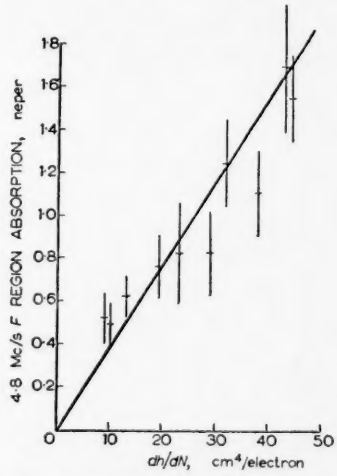


Fig. 6. The F-region absorption of 4·8 Mc/s waves plotted against dh/dN, the reciprocal of the electron density gradient at the reflection level.

variations. Actual values of dh/dN varied by factors of 2 and 3 from one day to the next even on magnetically quiet days.

In Figs. 5 and 6, the F-region absorptions of 4 and 4·8 Mc/s waves are plotted against the corresponding values of dh/dN obtained from the formulae given above. Each point represents the median of about 7 months' results. It is seen that the absorption is proportional to dh/dN. From the slopes of the lines it was deduced that the collision frequency at the level of reflection of 4 Mc/s waves was $(3{\cdot}6 \pm 0{\cdot}6) \times 10^3$/sec and at the level of reflection of 4·8 Mc/s waves was $(3{\cdot}0 \pm 0{\cdot}6) \times 10^3$/sec. The height of the former reflection varied from about 150 to 180 km and that of the latter from about 180 to 200 km.

6. Discussion of Results

6.1. *The E-region*

The collision frequency ν_0 was found to be $(1{\cdot}9 \pm 0{\cdot}1) \times 10^4$/sec. This is greater than the value of $(1{\cdot}3 \pm 0{\cdot}1) \times 10^4$/sec found by the writer (WHITEHEAD,

VOL. 16 1959

1959) from a study of the sudden changes in the virtual height associated with changes in the mean amplitude of the echo. Reasons for this discrepancy are being sought.

The height of the *E*-region showed no variation with the sunspot cycle. This would be expected on the Chapman theory which indicates that the height of the layer depends on the wavelength of the incoming radiation and not on its intensity. We may conclude then that the changes observed in the *E*-region during the sunspot cycle (such as the variation in the critical frequency) are solely due to changes in the intensity of the ionizing radiation.

6.2. *The D-region*

The increase in the *D*-region absorption with sunspot number may be due to either an increase in the overall electron density in the *D*-region, or a decrease in the heights of the low layers.

The decrease in the additional *D*-region absorption in winter with increasing sunspot number might arise through the ionizable constituent in the *D*-region being completely ionized by the normal component at sunspot maximum, so that only at sunspot minimum could additional ionization occur giving rise to the abnormal winter absorption.

6.3. *The F-region*

The collision frequency in the *F*-region is substantially constant over the range of heights 150 to 200 km. NICOLET (1953) gives the electronic collision frequency with neutral atoms as 200/sec at 160 km, much less than the total collision frequency found in the present analysis. Thus most collisions must take place with positive ions. However, if the temperature was the same at the level of reflections of 4 and 4·8 Mc/s waves, the collision frequency would increase with height because the electron density and consequently the density of positive ions increases with height. The collision frequency with positive ions decreases with increasing temperature (NICOLET, 1953) so that it appears that the temperature is increasing quite rapidly with height. Making use of Table 2 of NICOLET's paper, which gives numerical values for the collision frequency, we deduce that the temperature increases from about 170°K to about 240°K from the level of reflection of 4 Mc/s waves to that of 4·8 Mc/s waves, an increase in height of about 30 km.

The actual value of the collision frequency at this height should be compared with that found by previous workers. CHAPMAN (1956) using NICOLET's formula gives a value of ν at 150 km as 2×10^3/sec and at 200 km as 8×10^2/sec. FARMER and RATCLIFFE (1936) found a value of $1{\cdot}2 \times 10^3$/sec for the *F*-region though this value may have to be increased to about $2{\cdot}7 \times 10^3$/sec (WHITEHEAD, 1958).

It is thus seen that the collision frequency found for the *F*-region is rather greater than the value usually given.

REFERENCES

BOOKER H. G. 1935 *Proc. Roy. Soc.* A **150,** 267.

CHAPMAN S. 1931 *Proc. Phys. Soc. Lond.* **43,** 26 and 433.

CHAPMAN S. 1956 *Nuovo Cim.* **4,** (10) 1385.

DIEMINGER W.	1952	*J. Atmosph. Terr. Phys.* **2,** 340.
FARMER F. T. and RATCLIFFE J. A.	1936	*Proc. Phys. Soc. Lond.* **48,** 839.
GNANALINGHAM S. and WEEKES K.	1954	*The Physics of the Ionosphere* p. 63. Physical Society, London.
GREGORY J. B.	1956	*Aust. J. Phys.* **9,** 324.
MITRA S. K.	1952	*The Upper Atmosphere* p. 294. The Asiatic Society, London.
NICOLET M.	1953	*J. Atmosph. Terr. Phys.* **3,** 200.
PIGGOTT W. R.	1953	*Proc. Instn. Elect. Engrs.* Pt. III, **100,** 61.
SEDDON J. C.	1958	*J. Geophys. Res.* **63,** 209.
THOMAS J. O., HASELGROVE J. and ROBBINS A. R.	1957	*Tables of Ionospheric Electron Density, Slough* Cavendish Laboratory, Cambridge.
WHITEHEAD J. D.	1956	*J. Atmosph. Terr. Phys.* **9,** 276.
WHITEHEAD J. D.	1957	*J. Atmosph. Terr. Phys.* **10,** 12.
WHITEHEAD J. D.	1958	*J. Atmosph. Terr. Phys.* **12,** 150.
WHITEHEAD J. D.	1959	*J. Atmosph. Terr. Phys.* **16,** 99.

VOL.
16
1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 291 to 306. Pergamon Press Ltd. Printed in Northern Ireland

An investigation of the ionospheric *D*-region

J. A. FEJER and R. W. VICE
National Institute for Telecommunications Research, South African Council for Scientific and Industrial Research, Johannesburg, South Africa

(*Received* 20 *January* 1959)

Abstract—Two previously described methods of ionospheric investigation, the observation of weak echoes from the *D*-region (GARDNER and PAWSEY, 1953) and the measurement of ionospheric wave interaction (FEJER, 1955), are used to determine tentative profiles of electron density and collision frequency for the ionosphere below about 85 km.

1. INTRODUCTION

OUR knowledge of the ionosphere below about 85 km has in the past been obtained mainly from studies of the absorption of short waves and of the propagation of long waves. The interpretation of such studies is usually a difficult task.

Two more direct methods of investigation have been described recently. The first of these (GARDNER and PAWSEY, 1953) is based on the observation of weak echoes from heights below about 85 km. This method has the advantage that results are obtained at all times of the day, and it was used as the main tool in this investigation.

The second method (FEJER, 1955), the measurement of wave interaction, provides results only when the echo from the *E*-layer is at least 40–50 dB above the noise level, and it was used mainly as a check on the first method.

Previous applications of these two methods were of a preliminary nature and a need for more extensive observations with the aid of more powerful pulse transmitters existed. The present investigation tries to satisfy this need without claiming to have exhausted the possibilities inherent in either of the two methods. It goes, however, considerably further in some respects than the two original investigations.

Essential details of the equipment and of observational techniques are given in Section 2 of this paper.

The experimental results obtained from the observation of weak *D*-region echoes are described in Section 3. The results are interpreted in terms of electron densities and collision frequencies.

Data from the ionospheric wave interaction experiment are shown in Section 4 to be reasonably consistent with results obtained from the observation of weak echoes, in Section 3.

2. OBSERVATIONAL TECHNIQUES

Both types of observation described in the Introduction made use of the same two pulse transmitters operating simultaneously on different frequencies. The two pulse transmitters formed parts of a modified G L Mark II radar transmitter and used a common power supply with suitable decoupling.

One of the transmitters (used as the disturbing transmitter in the wave interaction experiment) operated on a frequency of 1·83 Mc/s with a pulse width of 30 μsec and a recurrence frequency of 75 c/s, delivering a power of about 70 kW

to the aerial feeders. About 20 kW was radiated by the circularly polarized aerial system which consisted of four half-wave dipoles 60 ft above the ground, arranged to form the sides of a square. Parallel dipoles were in phase with each other, and in quadrature with the perpendicular pair. The calculated radiation in the vertical direction was equivalent to an omnidirectionally radiated peak power of about 320 kW.

The other transmitter (used as the wanted transmitter in the wave interaction experiment) operated on a frequency of 2·63 Mc/s with a pulse width of 30 μsec and a recurrence frequency of 150 c/s. The peak radiated power was about 10 kW. The aerial system consisted of four half-wave dipoles 35 ft above the ground forming the sides of a square.

Observation of weak echoes from *D*-region was carried out using a common aerial system for both transmitting and receiving. Gas-filled spark gap switches were used in the conventional manner to switch over from reception to transmission, and vice versa. Provision was also made for the separate preset adjustment of the aerial polarization on transmission and reception, and for the simultaneous switching by relays of the sense of polarization on both transmission and reception.

Separate transmitting and receiving aerials had to be used on 2·63 Mc/s in the wave interaction experiment. One pair of parallel dipoles was used for transmission and the other pair for reception. The use of separate aerials was necessary because spark gap switches cause a small random variation in successive pulse amplitudes of the wanted transmitter, which masks the transferred modulation.

The method of detecting the transferred modulation in the wave interaction experiment has been described previously (FEJER, 1955). The amplitude modulation of the wanted echo at the recurrence frequency of the disturbing transmitter is detected by a suitable strobe, a selective amplifier and a phase sensitive detector. Provision is made for adjusting the time interval between the disturbing transmitter pulse and the wanted echo to values (expressed as heights) between 30 and 100 km.

A receiver of special design was used for the observation of weak echoes from the *D*-region. This receiver has a fast acting automatic gain control. The control voltage, which is approximately proportional to the logarithm of the instantaneous input voltage, is displayed after further amplification by a double beam cathode-ray tube using an *A*-type display. The other beam of the tube is used for the display of 10 km height calibration marks on the time base.

Photographs of the face of this tube are taken automatically every 10 sec. Two exposures, each of about $\frac{1}{4}$ sec duration, are taken in quick succession without moving the film. Between these two exposures the sense of spot deflexion caused by the receiver output voltage and the sense of rotation of the polarization for both transmission and reception are automatically reversed. During the 10 sec intervals between consecutive pairs of exposures the film is moved to the next frame.

Typical photographs of the display, taken on our two frequencies near midday are shown by Fig. 1. The top traces correspond to ordinary polarization, the bottom traces to extraordinary polarization of the aerial system.

Amplitude calibration is carried out by running the equipment with the transmitters inoperative and a standard signal generator weakly coupled to the receiver

VOL. 16 1959

OL.
16
959

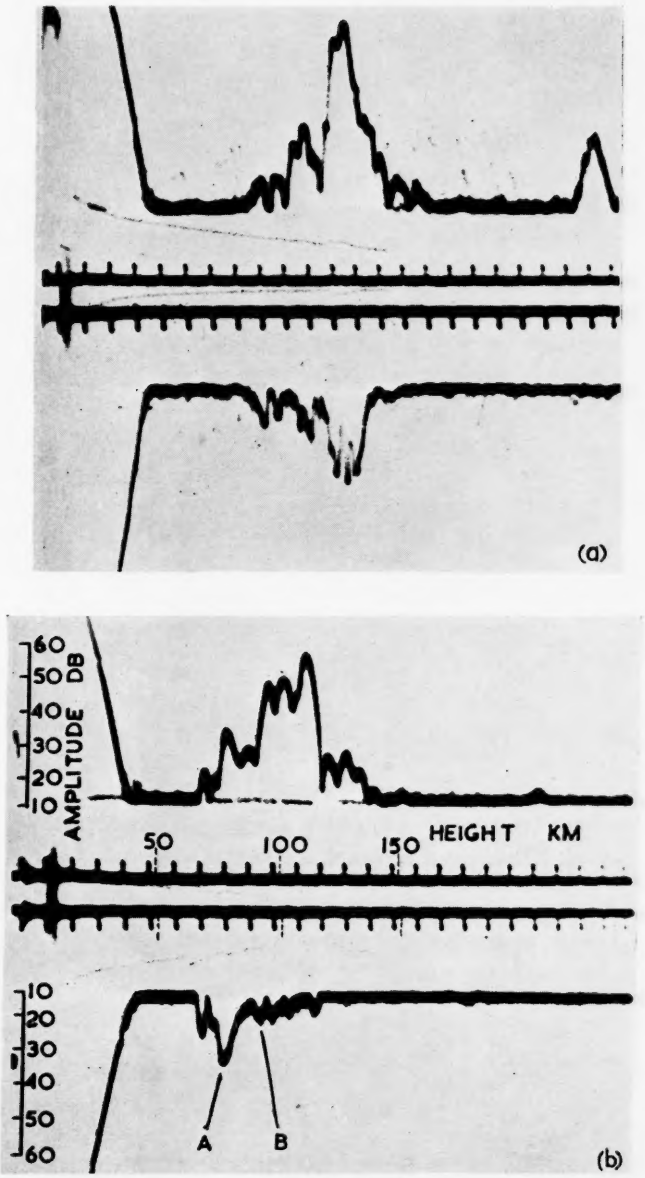


Fig. 1. Typical midday records taken at 1405 hours SAST on 17/4/1958. The indicated height calibration takes the receiver delay of 3 km into account. Record (a) was taken on a frequency of 2·63 Mc/s, record (b) on 1·83 Mc/s.

VOL.
16
1959

input. The signal level is changed in 10 dB steps between successive pairs of exposures.

The records obtained in this way were also used for the determination of the absorption of the main echo. Absolute amplitude calibration (in terms of equivalent reflection coefficient) was carried out in the usual manner by observing the amplitude of higher order echoes as well as of the main echo.

The large dynamic range of the logarithmic receiver, which was necessary for the recording of the amplitude of the main echo and of the weakest echoes from the *D*-region simultaneously, may have slightly reduced the accuracy of amplitude readings. Amplitudes could be read to the nearest decibel and this was regarded sufficient for most of the present observations.

Two different methods were used to set the polarization adjustments. In one of these the adjustments were set for minimum amplitude of the *E*-echo at midday with the polarization switch in the "extraordinary" position. It was assumed that the extraordinary component is completely absorbed at midday.

The other method could be used only at times when the difference in the virtual height of the two magneto-ionic components was sufficiently great to split the echo. Then, with the polarization switch again in the "extraordinary" position, the adjustments were set for minimum amplitude of the ordinary component of the echo as distinguished by its virtual height.

The two methods resulted in the same setting of the adjustments. The discrimination of the polarized aerial system against the unwanted component was about 40 dB which was quite sufficient for the present purpose, in view of the presence of other more serious limitations to the method, which will be described in the following sub-sections.

The limit in the discrimination against the unwanted component was mainly set by the sharpness of the dipole resonances, and the consequent frequency dependence of the aerial polarization, which was only correct in the centre of the pulse frequency spectrum. A lesser part was probably played by the deviation of the downcoming waves from the vertical direction.

3. Weak Echoes from the *D*-Region

3.1. *Introduction*

Previous observations of weak echoes from the *D*-region were described by Dieminger (1955), Gardner and Pawsey (1953), and Gregory (1956), among others. The techniques used by Gardner and Pawsey are followed in the present investigation with two modifications.

One of the modifications is the simultaneous switching of the polarizations on both transmission and reception, and the resulting better discrimination against the undesired magneto-ionic component.

The other more important modification in the method of recording and scaling will be discussed in Section 3.3.

It is known from rocket measurements (Jackson and Seddon, 1958) and from the results of a companion paper (Fejer and Vice, 1959) that the reflection point for frequencies near 2 Mc/s is not much above 90 km at midday at the latitude of

Johannesburg. The refractive index is therefore likely to differ substantially from unity at heights greater than about 85 km and this makes the interpretation of the observation of weak echoes difficult. In view of the low true height of the reflection point at midday it is also clear that echoes coming from observed heights of 90 to 95 km suffer considerable group retardation and come in reality from much lower heights. It is therefore reasonable to expect that useful results will not be obtained from the study of echoes coming from greater observed heights than about 85 km.

3.2. *Assumptions made in the interpretation*

The present method of interpretation is due to GARDNER and PAWSEY (1953). The amplitude ratio of the ordinary component to the extraordinary component is observed as a function of the height. From the ratio measured at the lowest height the collision frequency is estimated. The integrated differential absorption is then determined from the ratio observed at greater heights. The electron density may be calculated from the rate of change of the integrated differential absorption with height.

Two assumptions are made in this interpretation. First it is assumed that the refractive index does not differ greatly from unity. This assumption, already mentioned in the previous subsection, is certainly invalid in the neighbourhood of the point of reflection.

Even at lower heights where the mean refractive index is near unity, pockets of intense ionization of meteoric origin could exist. If the echoes were caused by meteoric ionization, they would be most intense in the early morning hours. Observations show, however, that the weak echoes below 85 km disappear at sunset and only reappear at sunrise. This fact alone rules out the possibility of meteoric origin of the echoes. The tendency, shown by Fig. 2, for echoes to occur at certain preferred heights is also difficult to reconcile with meteoric ionization.

VOL. 16 1959

It is unlikely that other ionizing processes could produce limited regions of very intense ionization. Such regions of intense ionization could also be formed by rearrangement of existing ionization but it is felt to be unlikely that regions thus formed could exceed the existing ionization by several orders of magnitude. Our first assumption is therefore likely to be fulfilled well below the point of reflection.

Secondly, it is assumed that the observed heights of the echoes may be interpreted as true heights. This assumption is valid only if the waves comprising the echo deviate very little from the vertical direction, and if at the same time the refractive index does not differ greatly from unity in the region between the scatterer and the observer. The latter condition is merely part of the first assumption. GARDNER and PAWSEY suggest that the frequently observed pronounced stratification is, to a certain extent, evidence for the almost vertical arrival of the echoes. This argument, unfortunately, does not decide whether the obliquely scattered energy is sufficient to cause weaker echoes at greater apparent heights than the echo showing strong stratification. For example, the echo marked B in Fig. 1(b), which shows a typical photograph of the screen at midday on a frequency of 1·83 Mc/s, could, on the evidence of the 1·83 Mc/s records alone, be equally well interpreted as coming obliquely from the same height as the echo marked A, or as coming from a vertical direction from a height equal to its observed height. It will

be shown in a later part of the paper, on the basis of records taken on both frequencies, that the echo marked B arrives obliquely, and that our second assumption is thus definitely invalid on the frequency of 1·83 Mc/s except for very low heights, although it remains valid up to greater heights on 2·63 Mc/s. The best ultimate test of the second assumption would be the measurement of the direction of arrival of the echoes.

3.3. *Method of recording and scaling*

The type of records used in the present investigation are shown by Fig. 1; they are *A*-type (deflexion) displays of the amplitude of both magneto-ionic components of the echoes on a suitably calibrated time base. There were twenty-four exposures available for every hour and for each of the two frequencies. These were examined for distinct echoes with both magneto-ionic components occurring at the same height. The amplitude of both magneto-ionic components and the height of twenty-four such echoes were scaled on each of the two frequencies for every hour.

This method of recording and scaling is undoubtedly more laborious than the photographic integration method used by GARDNER and PAWSEY, and one has to be satisfied with much smaller samples. An advantage of the present method is the simultaneous measurement of the amplitude of the two magneto-ionic components (the $\frac{1}{4}$ sec delay is negligible in this respect) and the elimination of the possibility of reading on the edges of echoes.

L.
6
59

3.4. *Observations*

The observations described in this section were carried out during the first 6 months of 1958 at Frankenwald (latitude 26° 4′S, longitude 28° 6′E) near Johannesburg. Observations were taken between the hours of 0400 and 2000 SAST on about 8 days of each month which included the Regular World Days of the International Geophysical Year. Only the first quarter of each hour was available for the observation of weak echoes from the *D*-region and for absorption measurements and the forty-eight exposures (twenty-four on each frequency) mentioned in the previous subsection were taken during this interval after the amplitude calibration.

Before considering the observed amplitude ratios of the magneto-ionic components on which most of the results of the present investigation are based, it is interesting to consider the absolute values of echo amplitude expressed in terms of the equivalent reflection coefficient. One should keep in mind that only distinct echoes were scaled and no entry was made at times when there was no echo. Our measured reflection coefficient is therefore likely to be an overestimate.

Another point to be kept in mind in the interpretation of measured equivalent reflection coefficients is the tendency of echoes to originate from certain preferred heights. This is shown clearly by Fig. 2 which indicates the relative frequency with which echoes were received at noon in April 1958 from different 2 km wide height intervals. The small number of echoes received from about 75 and 85 km is easily noticed. The histogram of Fig. 2 is similar to other histograms obtained for different times of the day and the year.

The tendency to attribute the scarcity of echoes from certain heights to low electron densities should be resisted until an independent proof is obtained that

there is a correlation between the frequency of occurrence of the echoes and the electron density.

The values obtained for the equivalent reflection coefficient should be considered with the above reservations. In the determination of this reflection coefficient the observed values of the amplitude were corrected for the absorption integrated up to the height of the echo. After this, correction values of the order of 3×10^{-5} for a height of 75 km and 10^{-4} for a height of 85 km were obtained for the frequency of 1·83 Mc/s and about half these values for the higher frequency. It is interesting

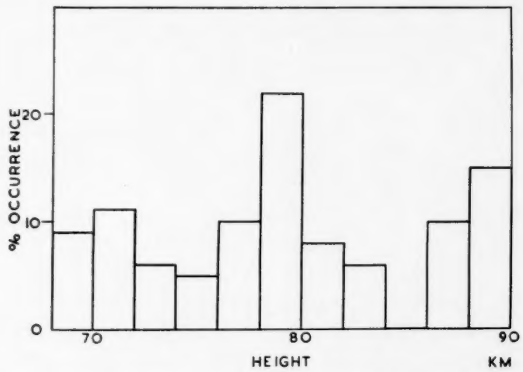


Fig. 2. Histogram showing percentage of echoes obtained from different height intervals in April between 1100 hours and 1315 hours SAST.

VOL. 16 1959

to compare the ratio of these two values (about 3 : 1) with the ratio of the electron densities, for which a value of 30 : 1 is obtained from results derived in later parts of this paper. The discrepancy between these two ratios is a strong argument against associating the equivalent reflection coefficient with the electron density. It appears that the irregularities are relatively more intense at lower heights. In the same way the gaps in the frequency of occurrence of the echoes shown by Fig. 2 could be caused by a less irregular distribution of ionization at the heights concerned rather than by less ionization.

No special study of the fading characteristics of the weak echoes was carried out. The echoes appeared to fade faster than the main echo; the time between successive fades was of the order of 5–15 sec at all observed heights below the main echo.

Strong correlation between the fading of the magneto-ionic components was observed at the lowest heights and the amplitude ratio of the two components was almost constant. The correlation coefficient decreases with height from values of the order of 0·75 at 70 km to 0·4 at 85 km.

The scaled data were obtained in the form of observed values of the amplitude ratio of the magneto-ionic components of echoes together with the observed values of height of the same echoes. Before any use could be made of these data they had to be sorted according to height. Observations for each hour of the day and for each month were sorted into 5 km wide height intervals. In addition the combined observations for 1100–1315 SAST were sorted into 2 km wide height intervals for each month.

Fig. 3 shows the median value of the amplitude ratio (of the magneto-ionic

components) as a function of the hour of the day for the different 5 km wide height intervals, computed for the combined data of January, February and March 1958. The curves for 1·83 Mc/s shown by Fig. 3(b) have maxima at midday only for the three lowest height intervals. The curves labelled 79–84 km and 84–89 km on the other hand show definite midday minima with maxima in the early morning and late afternoon.

The reason for this midday minimum can be understood by looking at the echo pattern of Fig. 1b. If it is assumed that the basic assumptions made in Section 3.2

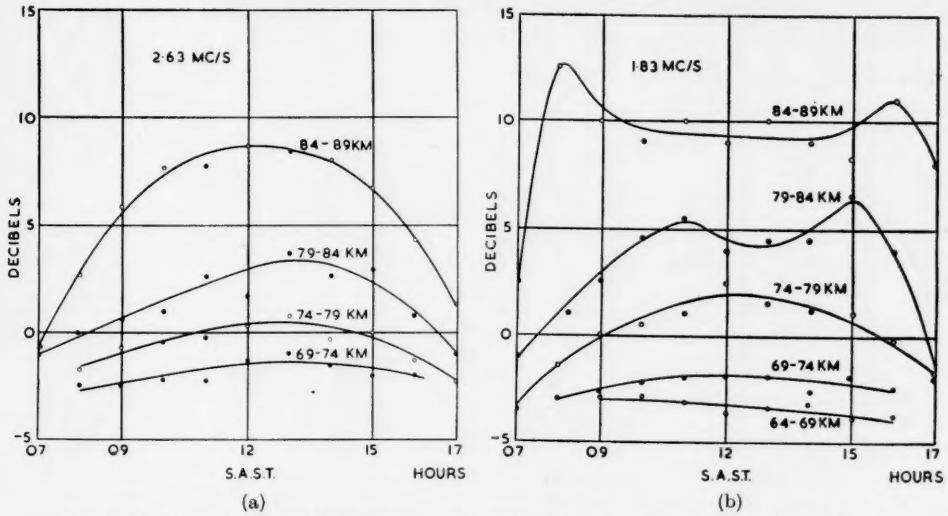
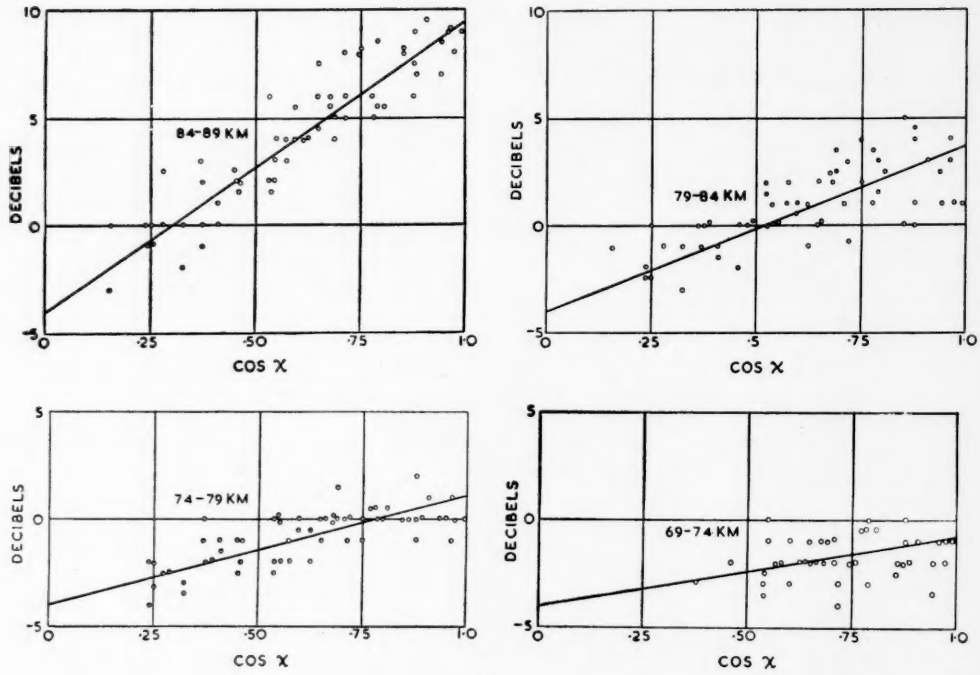


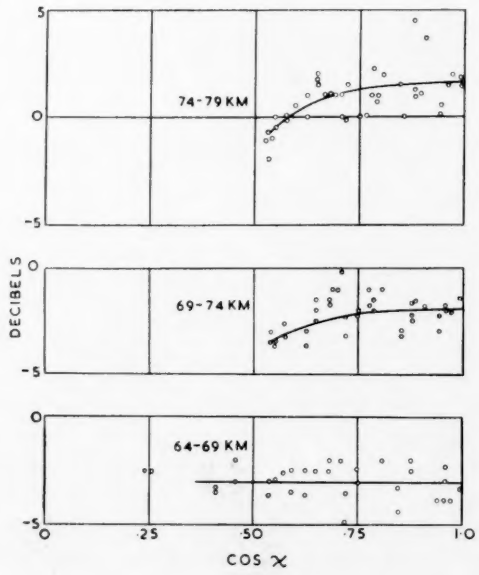
Fig. 3. Median values of the hourly amplitude ratio (ordinary component to extraordinary component) for the combined data of January, February and March 1958 as a function of the time of the day with the height interval as the parameter. The (a) curves refer to 2·63 Mc/s, the (b) curves to 183 Mc/s.

are valid on 2·63 Mc/s and the resulting electron densities are calculated (as in the later parts of this paper) then it can be shown that at the height of the echo marked B, on the trace taken with extraordinary polarization of the aerial system, the differential absorption is so high that the extraordinary component of the echo shown at the same height as echo B, on the trace taken with ordinary polarization. ought to be well below the noise level. Hence the true height of the echo marked B cannot be its observed height. The most reasonable assumption is that the echo marked B is the extraordinary component of an echo scattered obliquely from the lower observed height of the stronger echo marked A. (The discrimination of the aerial system against the ordinary component is sufficient even at this oblique angle to discount the possibility that the echo marked B may be the ordinary component.)

Thus only conclusions about the scattering polar diagram of the irregularities, but not about electron densities, may be drawn from records taken on the frequency of 1·83 Mc/s near midday for heights greater than about 80 km. One would agree with GARDNER and PAWSEY that the bulk of the received radiation has a semicone angle of about 15°, indicating the presence of flat horizontal irregularities, with the



(a)



(b)

Fig. 4. Monthly median values of the hourly amplitude ratio (ordinary component to extraordinary component) plotted as a function of cos χ. The (a) graphs refer to 2·63 Mc/s, the (b) graphs to 1·83 Mc/s.

VOL. 16 1959

reservation that a small part of the back-scattered energy arrives at angles as much as 30–50° removed from the vertical.

In view of this information about the scattering polar diagram it is reasonable to assume that the second assumption of Section 3.2 is valid if the amplitude of the extraordinary component does not decrease with height. This criterion is generally satisfied up to a height of about 86 km on 2·63 Mc/s and up to about 76 km on 1·83 Mc/s near midday.

It is convenient to plot the measured amplitude ratios of the magneto-ionic components against the solar zenith angle χ rather than against local time. In Fig. 4 the results of sorting into 5 km height intervals as previously described are displayed in this way for height intervals of 69–74, 74–79, 79–84 and 84–89 km for the frequency of 2·63 Mc/s and for height intervals of 64–69, 69–74 and 74–79 km for the frequency of 1·83 Mc/s. At the frequency of 2·63 Mc/s the trend of the points is reasonably well expressed by the straight lines drawn in Fig. 4; curves were drawn to approximate the points at 1·83 Mc/s.

An upper limit to the collision frequency at the lowest height of observation (about 66 km) is obtained by assuming that no differential absorption occurs below this height, and calculating the collision frequency from the measured amplitude ratio of the magneto-ionic components at this height. From Fig. 4 this ratio is about 1·4–1·6 on 1·83 Mc/s resulting (using formulae of the magneto-ionic theory as given by GARDNER and PAWSEY) in a collision frequency of about 10^7–$1{\cdot}4 \times 10^7$. As this value is an upper limit, it rules out the value of $3{\cdot}5 \times 10^7$ resulting from the collision frequency profile suggested by NICOLET (1953). The value of 9×10^6 following from the collision frequency profile suggested by CROMPTON, HUXLEY and SUTTON (1953) on the basis of their laboratory measurements, is just a little below the upper limit set by data. This profile shown by Fig. 9 was used in all subsequent computations.

Fig. 5 shows the computed ratio of the reflection coefficients for the two magneto-ionic components (for weakly ionized irregularities) and the computed differential absorption coefficient for unit electron density as functions of the height for the adopted collision frequency profile.

With the aid of Fig. 5 it may be ascertained that the straight lines fitted to the points of Fig. 4 cut the $\cos \chi = 0$ axis at values very nearly equal to the theoretical ratio of the reflection coefficients of the magneto-ionic compounds, which for a frequency of 2·63 Mc/s is near 5 dB for the height intervals of Fig. 4. This shows that, at least for heights above 71 km, the integrated differential absorption, and thus the electron density, is very nearly proportional to $\cos \chi$.

If the data are to be interpreted in terms of electron density, it is necessary to determine the rate of change of the integrated differential absorption with height. The sorting into 5 km wide intervals does not appear fine enough for this purpose. Therefore the combined 2·63 Mc/s data of the three hours near midday, which correspond to almost identical solar zenith angles, were sorted into 2 km wide height intervals for each month. The median values of the ratio of the magneto-ionic components for these groups are shown by the points of Fig. 6 for January as a function of the mean height of the group. Smooth curves were drawn to approximate these points as in Fig. 6, and the values of these curves for certain selected

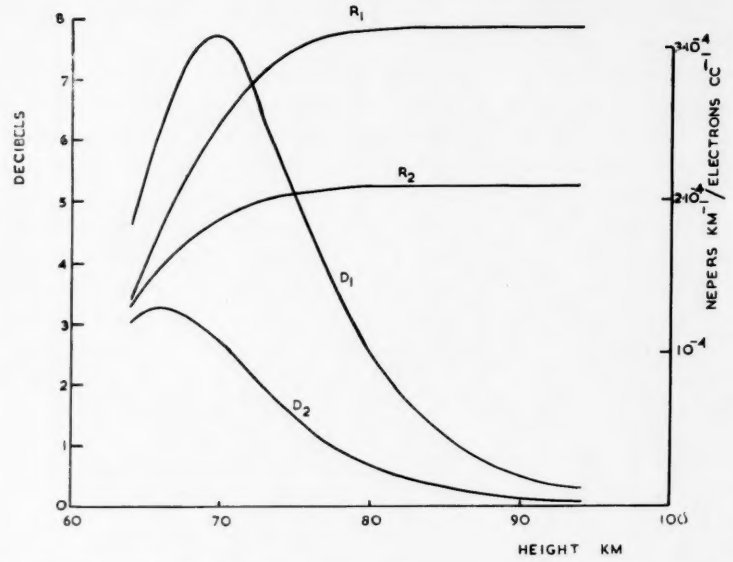


Fig. 5. The computed ratio of reflection coefficients (extraordinary component to ordinary component) in decibels for 1·83 Mc/s (curve R_1) and 2·63 Mc/s (curve R_2) and the computed differential absorption per unit electron density for 1·83 Mc/s (curve D_1) and 2·63 Mc/s (curve D_2) as functions of the height.

VOL. 16 1959

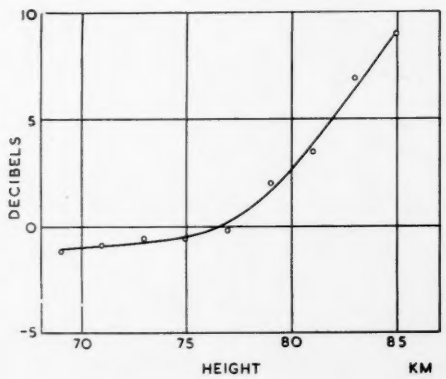


Fig. 6. Monthly median value of the amplitude ratio (ordinary component to extraordinary component) at midday (between 1100 hours and 1315 hours SAST) for 2 km wide height intervals as a function of the mean height of the interval.

heights are then plotted against cos χ in Fig. 7. Each set of points of Fig. 7 corresponding to the same height was then approximated by a straight line cutting the cos $\chi = 0$ axis at -5 dB which is near enough to the theoretical ratio given by Fig. 5 for any of the heights used in Fig. 7. Based on the intersection of these straight lines with the cos $\chi = 1$ line, a curve of integrated differential absorption as a function of the height was drawn. This curve could have been based on the straight lines of Fig. 4 without changing the result substantially; no significant

finer detail was obtained owing to sorting into narrower height intervals. The intersections of the curves of Fig. 4(b) with the $\cos \chi = 1$ lines were in fact used at the frequency of 1·83 Mc/s.

From the rate of change of integrated differential absorption, as a function of the height, the electron density was calculated using the theoretical curves of Fig. 5 for the differential absorption coefficient for unit electron density. The result is shown by Fig. 8. Curve *A* represents the electron densities derived from the 1·83 Mc/s data while curve *B* shows the electron densities obtained from the

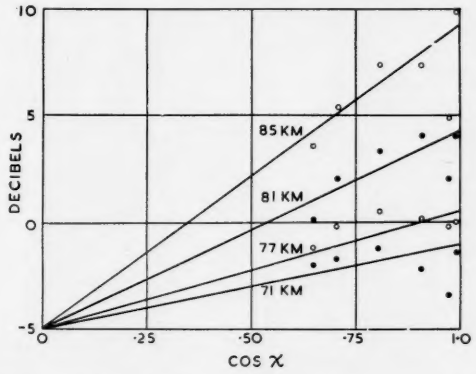


Fig. 7. Monthly smoothed median values of the amplitude ratio (ordinary component to extraordinary component) based on curves of the type shown by Fig. 6, as a function of $\cos \chi$ with the height as parameter.

2·63 Mc/s data. Curve *C* shows the results obtained from rocket measurements (JACKSON and SEDDON, 1958) for greater heights. In Fig. 9 curves *A*, *B* and *C* were combined into one smoothed curve of electron density as a function of the height; the adopted collison frequencies are also shown. Unfortunately, the present data furnish no information on the electron density distribution below 68 km, apart from the fact that very weak echoes are still observed regularly at heights of 60 km near midday. In the absence of quantitative results the electron density curve of Fig. 9 was not drawn below 68 km.

It must be emphasized that the electron density curves *A* and *B* of Fig. 8 are calculated from the rate of change of the experimentally determined integrated differential absorption with height. The smaller one makes the height difference in the determination of this rate of change, the greater the error becomes. There is thus little hope of reliably determining the finer details of the electron density profile by the present method.

The electron densities given by Fig. 9 for $\cos \chi = 1$ and assumed proportional to $\cos \chi$, representing the results of this investigation, are difficult to explain by a known mechanism of layer formation. In view of the inability of the method of investigation to detect finer structure, the true electron density profile may be more complicated than the smooth curve of Fig. 9. The tendency of the echoes, shown by Fig. 2, to come from certain preferred heights may be an indication (though it is not evidence) of three distinct ionospheric layers below the *E*-layer in day-time.

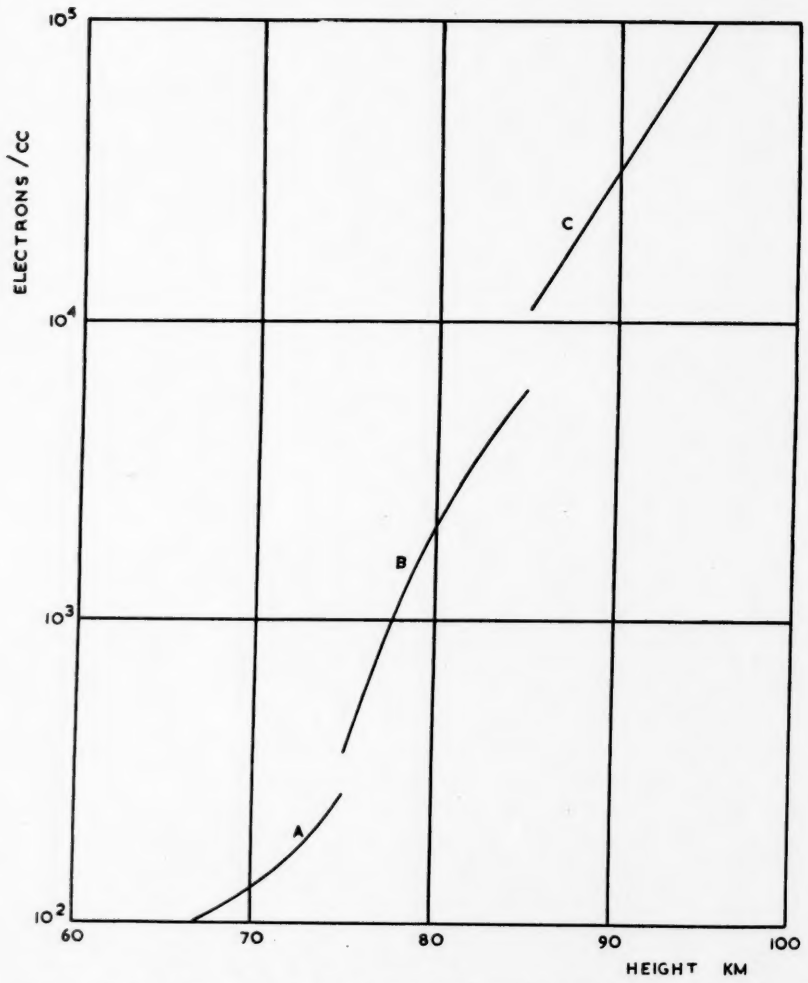


Fig. 8. Electron density for $\cos \chi = 1$ as a function of the height. Curve A is based on present data for 1·83 Mc/s, curve B on present data for 2·63 Mc/s, while curve C represents the results of rocket measurements.

Further more definite indication of at least one separate layer below 75 km (also pointed out by Gardner and Pawsey) is the decreased slope of the electron density profile below 75 km shown by Fig. 8 and Fig. 9.

4. Wave Interaction in the *D*-Region

The present observations on the interaction of pulsed radio waves in the *D*-region were virtually a repetition of the previous preliminary work (Fejer, 1955) using more transmitter power, higher recurrence frequency and shorter pulses. Details of the equipment are given in Section 2.

It will be recalled that in this experiment both the disturbing and the wanted

VOL. 16 1959

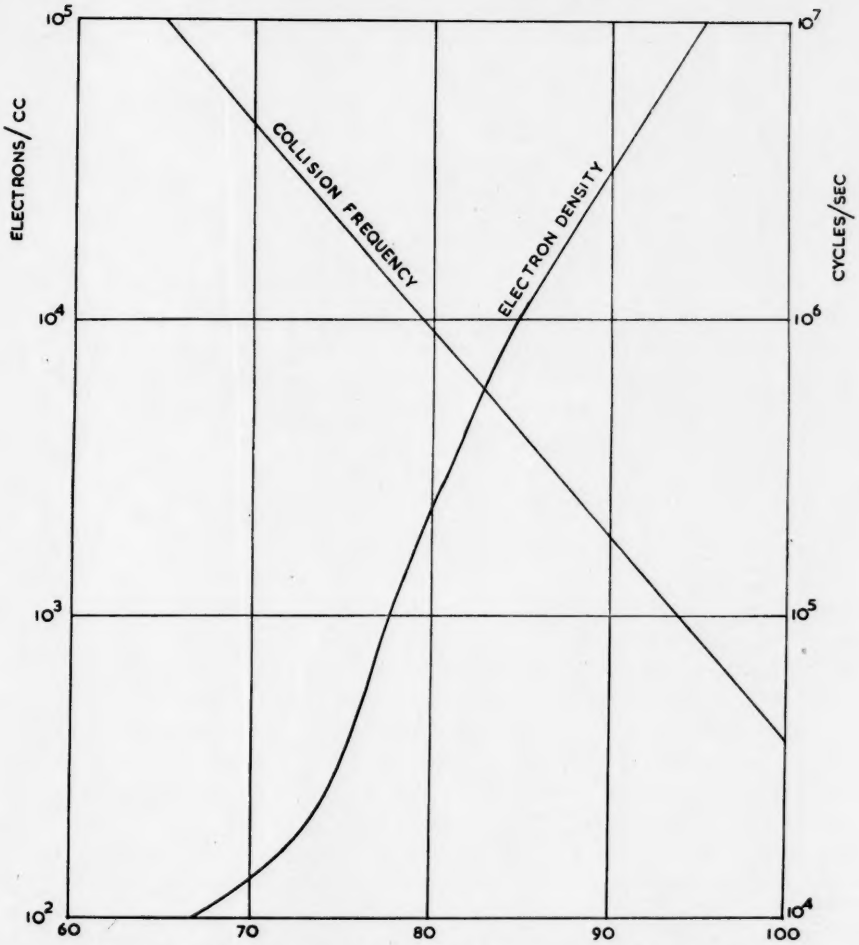


Fig. 9. Collision frequency and electron density as functions of the height.

transmitters are pulsed and the time elapsed between the transmission of the disturbing pulse and the arrival of the wanted echo is varied. This time delay t_0 is conveniently expressed in terms of the height $h_0 = ct_0/2$ at which the upgoing disturbing pulse and the downcoming wanted pulse meet in the ionosphere. The disturbing transmitter operates at half the recurrence frequency of the wanted transmitter and the transferred modulation of the received echo at half the recurrence frequency is detected by a selective amplifier and a phase sensitive detector.

A section of the record of transferred modulation for different values of h_0 is shown by Fig. 10. The records were considerably better than those obtained in the preliminary investigation but there was still very little reserve of signal-to-noise ratio. In view of this difficulty the method was used mainly as a check on the model of Fig. 9 obtained in Section 3.

Observations were carried out during ten days between the 17 July and the 27 August, 1957. Days with relatively low absorption were selected because the method fails if the echo is not at least 40–50 dB above the noise level.

The height h_0 was set in 10 km steps. No transferred modulation was observed when h_0 was set to values of 30, 40, 50 and 60 km. The measured transferred modulation, using extraordinary polarization of the disturbing wave, is shown for values of h_0 of 70, 80, 90 and 100 km by the points of Figs. 11(a)–(d). Each point corresponds to a record of about 5 min duration. The curves are calculated from

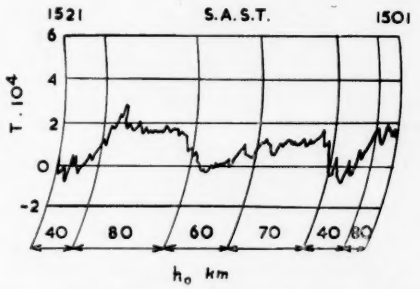


Fig. 10. Sample of record of transferred modulation as a function of the time (increasing towards the left). The set values of the quantity h_0 are indicated at the bottom.

the model of Fig. 9 for $\cos \chi = 1$, assuming proportionality between $\cos \chi$ and the electron density. Equation (8) of FEJER (1955) was used in this calculation.

Examination of Fig. 11 shows that although there is a large amount of scatter, the diurnal trends indicated by the points and the curves are similar. The order of the observed transferred modulation agrees reasonably well with the calculated curves except in the case of $h_0 = 70$ km, where the points are rather below the curve indicating that the electron density resulting from the wave interaction experiment for heights just below 70 km is lower by a factor of about 1·5. This is not a serious discrepancy even without mentioning that the points in Fig. 11 are based on measurements in 1957 while the curves are based on measurements in 1958.

The curves of Fig. 11 were calculated using the value of $G = 0{\cdot}01$ for the energy-loss coefficient which again fitted the data best. The present results thus again confirm approximately this value of G, and the collision frequencies obtained in the preliminary investigation.

During most of the present observations extraordinary polarization of the disturbing wave was used, as this leads to a larger and more easily measured transferred modulation and as it gives information about electron densities at a lower height. A few observations were made using ordinary polarization of the disturbing wave. The results agree in order of magnitude with curves derived from the model of Fig. 9, but they are not sufficiently numerous to justify their inclusion here.

Summing up, it may be stated that the present observations of wave interaction have gone a little further than the previous preliminary measurements. Despite the scatter of the individual points in Fig. 11 a trend of diurnal variation can be discerned. Still larger transmitter power and a receiving site more free of noise would probably result in more reliable individual records.

VOL. 16 1959

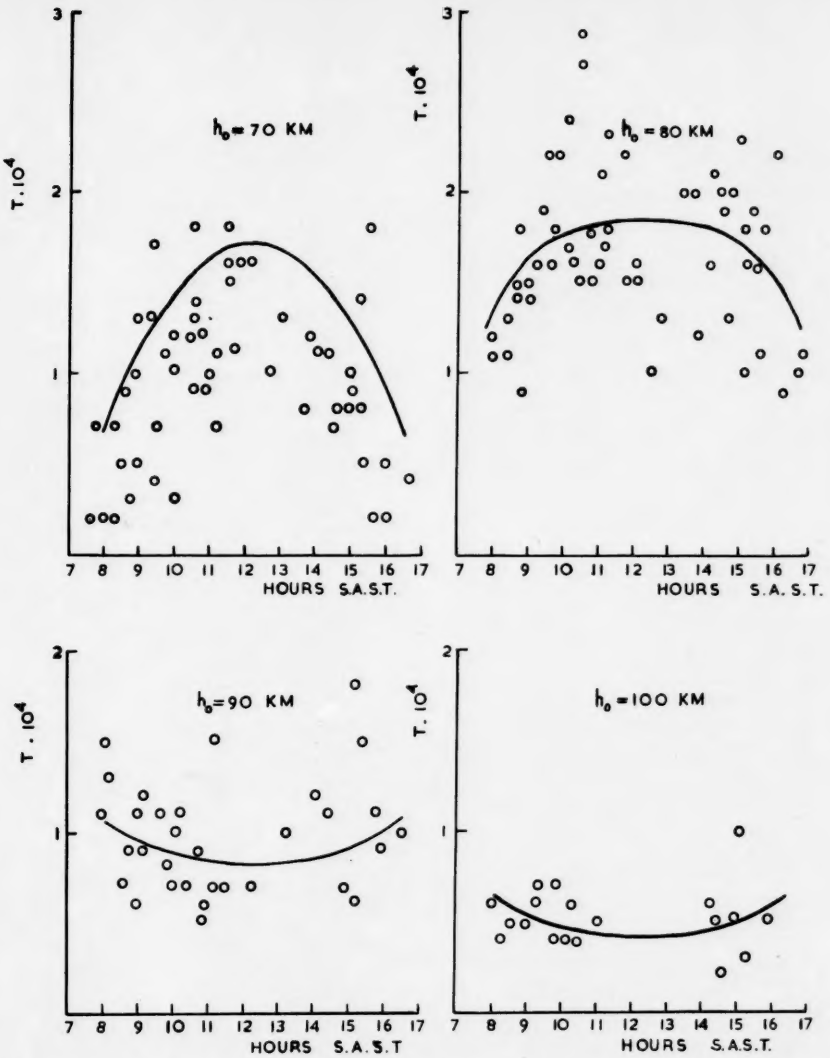


Fig. 11. The points indicate the measured mean transferred modulation T for samples of about 5 min duration as a function of the time of the day. The curves were computed from the model of Fig. 9.

5. Conclusions

A model of the lower ionosphere between the heights of about 68 and 85 km is obtained in this paper. The model is based on the rate of change of the measured integrated differential absorption with height, and therefore shows only the smoothed variations with height. Other methods will have to be used to ascertain whether there are any rapid changes of electron density with height not detected by the present method which may indicate the presence of separate ionospheric layers.

Some evidence for the existence of such a separate layer below 75 km is obtained from the present observations.

An interesting property of the model is the proportionality of the electron density to the cosine of the solar zenith angle, at least for heights between 70 and 85 km.

Very little absorption is suffered by a high frequency radio wave while passing through the region below 85 km. It accounts only for about one-sixth of the total absorption at 1·83 Mc/s at midday and for an even smaller fraction at higher frequencies.

Acknowledgements—The authors would like to thank Messrs. F. W. R. Potgieter and K. G. Kreft for assistance with the observations and Mr. A. Boyd for scaling most of the data. Thanks are also due to the South African Council for Scientific and Industrial Research for permission to publish this paper. This work formed part of the South African programme for the International Geophysical Year.

References

Crompton R. W., Huxley L. G. H. and Sutton D. J.	1953	*Proc. Roy. Soc.* A **218,** 507.
Dieminger W.	1955	*The Physics of the Ionosphere* p. 53. Physical Society, London.
Fejer J. A.	1955	*J. Atmosph. Terr. Phys.* **7,** 322.
Fejer J. A. and Vice R. W.	1959	*J. Atmosph. Terr. Phys.* **16,** 307.
Gardner F. F. and Pawsey J. L.	1953	*J. Atmosph. Terr. Phys.* **3,** 321.
Gregory J. B.	1956	*Aust. J. Phys.* **9,** 324.
Jackson J. E. and Seddon J. C.	1958	*J. Geophys. Res.* **63,** 197.
Nicolet M.	1953	*J. Atmosph. Terr. Phys.* **3,** 200.

VOL. 16 1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 307 to 317. Pergamon Press Ltd. Printed in Northern Ireland

The use of full-wave solutions in the interpretation of ionospheric absorption measurements

J. A. FEJER and R. W. VICE
National Institute for Telecommunications Research, South African Council for Scientific and Industrial Research, Johannesburg, South Africa

(*Received* 20 *January* 1959)

Abstract—A numerical method is developed for solving the equations of ionospheric wave propagation for vertical incidence. The method, which neglects coupling, is used for the calculation of ionospheric absorption.

The absorption is first calculated in the usual manner by using the Appleton–Hartree formula and integrating up to the theoretical height of reflection. A "full-wave" correction to this result is then obtained from a step-by-step integration of the wave equation. The thickness of the region, for which this integration has to be carried out, is only a few free space wavelengths, and therefore the additional time required for the calculation of the "full-wave" correction is relatively short. Numerical examples show that for frequencies of about 2 Mc/s this correction cannot be neglected.

The method is applied in a preliminary interpretation of observed values of absorption. It is concluded that more than half of the observed absorption occurs less than 2 km below the theoretical height of reflection.

1. INTRODUCTION

THE larger part of the day-time absorption of high frequency radio waves was until recently (MITRA, 1952) thought to occur in a separate ionospheric layer situated below the *E*-layer. Some doubt was cast on this assumption by BEYNON and DAVIES (1955) who showed that ionospheric absorption at 2 Mc/s can be better explained by assuming that it occurs in the process of reflection by a simple Chapman layer, the *E*-layer. The collision frequencies resulting from their investigations agree with those proposed by NICOLET (1953).

More recently electron density measurements with the aid of rockets (JACKSON and SEDDON, 1958) showed that the shape of the *E*-layer differs considerably from that of a Chapman layer. At the same time extrapolated laboratory measurements by CROMPTON *et al.* (1953), observation of weak echoes from *D*-region by GARDNER and PAWSEY (1953) and an ionospheric interaction experiment (FEJER, 1955) all result in collision frequencies which are smaller by a factor of about 4 than those suggested by NICOLET.

The purpose of the present paper is the comparison of the absorption calculated for an ionospheric model based on these more recent investigations with measured values.

Absorption is usually calculated by integrating the absorption coefficient (the imaginary part of the Appleton–Hartree refractive index) with respect to height for both the upgoing and the downcoming wave. The problem in such calculations is the upper limit in height at which the integration should be stopped. This is generally taken (MILLINGTON, 1938) as the height for which the electron density reaches its "critical" value neglecting collisions. In the present paper this height is called the theoretical height of reflection.

It is difficult to justify this choice of the upper limit theoretically, if the minimum in the absolute value of the refractive index is not very much smaller

than unity. In the present paper the full-wave solution is obtained and is used to determine the correct value of absorption.

The method of obtaining these full wave solutions is described in Section 2 of this paper. In Section 3 the method is applied to the calculation of absorption for a definite ionospheric model. The implications of the results of this calculation are discussed in Section 4.

2. METHOD OF CALCULATION

FÖRSTERLING (1942) has shown that in the case of vertical incidence the propagation of electromagnetic waves in an ionized medium is governed by the so called coupled wave equations. When coupling is neglected, the two magneto-ionic components are propagated independently. The propagation of each of these components is then expressed by the wave equation

$$\Pi'' + n^2\Pi = 0 \tag{1}$$

where n is the complex refractive index given by the Appleton–Hartree formula taken with the appropriate sign of the square root, and where the independent variable is the distance s measured positive downward in units of $\lambda/2\pi$, λ being the wavelength in free space. The quantity Π in equation (1) is a linear combination of the electric field components at right angles to the direction of propagation. The coefficients in this linear combination are functions of only the height in a horizontally stratified ionosphere.

It is well known that if n varies slowly with the distance s, the so called W.K.B. solutions (JEFFREYS, 1924) satisfy equation (1). One of these solutions $n^{-\frac{1}{2}} \exp\left(-i\int_{s_0}^{s} n\,ds\right)$ represents a downgoing wave while the other $n^{-\frac{1}{2}} \exp\left(i\int_{s_0}^{s} n\,ds\right)$ represents an upgoing wave (omitting the time factor $\exp(i\omega t)$). The full solution of equation (1) may then be written as the linear combination of these two waves

$$\Pi = An^{-\frac{1}{2}} \exp\left(-i\int_{s_0}^{s} n\,ds\right) + Bn^{-\frac{1}{2}} \exp\left(i\int_{s_0}^{s} n\,ds\right) \tag{2}$$

where A and B are constants. Differentiating results in

$$\Pi' = An^{-\frac{1}{2}}(-in - n'/2n) \exp\left(-i\int_{s_0}^{s} n\,ds\right) + Bn^{-\frac{1}{2}}(in - n'/2n) \exp\left(i\int_{s_0}^{s} n\,ds\right). \tag{3}$$

Dividing equation (3) by equation (2), and writing $A/B = R$, one obtains

$$\Pi'/\Pi = \left[R(-in - n'/2n) \exp\left(-2i\int_{s_0}^{s} n\,ds\right) + (in - n'/2n)\right] \Big/ \left[R \exp\left(-2i\int_{s_0}^{s} n\,ds\right) + 1\right]. \tag{4}$$

If the W.K.B. solutions were valid, R would be a constant. For this reason it is attractive to choose R as defined by equation (4) rather than Π as the dependent variable in equation (1). It is then necessary only to solve the differential equation for R as defined by equations (1) and (4) for a thin region

VOL. 16 1959

in the neighbourhood of the theoretical height of reflection, where the W.K.B. solution is known to be invalid. Towards the bottom of this region R approaches a constant value $R_0 = \lim_{s \to \infty} R$. Having determined the value of R_0, the solution valid for regions below the immediate vicinity of the theoretical height of reflection is given by equation (2) as

$$\Pi = B\left[R_0 n^{-\frac{1}{2}} \exp\left(-i\int_{s_0}^{s} n\, ds\right) + n^{-\frac{1}{2}} \exp\left(i\int_{s_0}^{s} n\, ds\right)\right]. \tag{5}$$

The reflection coefficient ρ at any point in these regions is equal to the ratio of downcoming to upgoing wave amplitude and with the aid of equation (5) it is given by

$$\rho = R_0 \exp\left(-2i\int_{s_0}^{s} n\, ds\right). \tag{6}$$

The upper limit of integration in (6) is determined by the height for which the reflection coefficient is required. More will be said about the lower limit later.

Although the substitution defined by (4) appears rather complicated, the resulting differential equation for R is relatively simple. The arithmetic of deriving the differential equation for R is simplified by using the intermediate variable

$$r = R \exp\left(-2i\int_{s_0}^{s} n\, ds\right) \tag{7}$$

and writing for (4)

$$\Pi'/\Pi = [r(-in - n'/2n) + (in - n'/2n)]/(r + 1). \tag{8}$$

Differentiating (8) results in

$$\Pi''/\Pi - \Pi'^2/\Pi^2 = d/ds\{[r(-in - n'/2n) + (in - n'/2n)]/(r + 1)\}. \tag{9}$$

Substituting Π''/Π from the differential equation (1) and Π'/Π from (8), equation (9) becomes a differential equation for r which after considerable simplification becomes

$$r' = i[n''/(4n^2) - 3n'^2/(8n^3)](r + 1)^2 - 2nri \tag{10}$$

or writing

$$N = (2n)^{-\frac{1}{2}} \tag{11}$$

equation (10) assumes the form

$$r' = -iNN''(r + 1)^2 - 2nri. \tag{12}$$

Substituting (7) into equation (12) one obtains the required differential equation for R

$$R' = -iNN''\left[R^2 \exp\left(-2i\int_{s_0}^{s} n\, ds\right) + 2R + \exp\left(2i\int_{s_0}^{s} n\, ds\right)\right]. \tag{13}$$

This differential equation with N defined by (11) is in a very suitable form for the numerical solution of the wave equation (1). Numerical integration has to be carried out only in the immediate neighbourhood of the point of reflection as the following reasoning shows.

A few free space wavelengths above the point of reflection the refractive index does not vary too fast and W.K.B. solutions are therefore approximately valid. Should a downcoming wave exist at this height, it would have an impossibly large amplitude only a little above this height, in view of the high absorption coefficient. It is therefore reasonable to assume a pure upgoing wave or $R = 0$ at this height. Step-by-step integration of equation (13) is then carried out downward (in the direction of increasing s). When a height, a few free space wavelengths below the theoretical height of reflection, is reached in the course of this integration, the W.K.B. solution again becomes a good approximation and thus R approaches a constant value $R_0 = \lim_{s \to \infty} R$. The reflection coefficient at still lower heights is then given by equation (6).

It is clear from the foregoing that the result will not be influenced by the choice of s_0, the lower limit of integration in (2). It is convenient to choose the theoretical height of reflection as the lower limit s_0 of integration because, with that choice, R_0 shows the error which would result by merely integrating the absorption coefficient with respect to height up to the theoretical height of reflection.

Equation (13) is similar to a differential equation derived by SCHELKUNOFF (1951) for a differently defined variable R. SCHELKUNOFF's equation (with which the present writers became familiar in the form given by GIBBONS (1955) after developing the present method) has the advantage of containing only the first and not the second derivative of the refractive index. One advantage of equation (13) is its use of the W.K.B. solution rather than a wave of constant amplitude in the definition of R, and the consequent possibility of restricting the integration to a much narrower range of heights. A second and more important advantage is the fact that while SCHELKUNOFF's R continues to vary in phase even in a slowly varying medium, with the present definition R remains constant in phase and amplitude, thus avoiding accumulated errors in the integration process. As a result of such errors the tip of a rotating vector moves outward on a spiral instead of moving on a circle. A third and less tangible advantage of the present method is the intuitive picture of the reflection process obtained directly from the results as displayed by Fig. 1 for the example to follow.

Details of the method are best illustrated by a numerical example. Let the electron density be a linear function of height. Using APPLETON's (1932) notation for the quantities x, y, z occurring in the formula for the refractive index, the value of x is unity for the theoretical height of reflection of the ordinary wave. Let $x = 0{\cdot}9$ for a height 0·52 km below the theoretical height of reflection. Let the frequency be 1·83 Mc/s so that 0·52 km corresponds to 20 units of $\lambda/2\pi$. Further let $y_L = 0{\cdot}417$, $y_T = 0{\cdot}215$ (values above Johannesburg) and $z = 0{\cdot}025$.

The integration was carried out from $x = 1{\cdot}03$ to $x = 0{\cdot}9$. The refractive index n and the quantity N were first calculated for length intervals of 0·052 km (2 units of $\lambda/2\pi$). The quantity N'' was then obtained by numerical differentiation. A desk calculating machine with six figure accuracy was used up to this point to avoid large errors in the numerical differentiation. The numerical integration was carried out to slide rule accuracy using intervals of 0·4 units of $\lambda/2\pi$ (about 0·01 km). The

VOL. 16 1959

complex values of R obtained in this way are shown by Fig. 1 which is very similar to a Cornu spiral. It was assumed that R tends towards $R_0 = -0{\cdot}18 + j0{\cdot}008$ which may be expressed as 15 dB additional attenuation due to the "full-wave correction". The absorption calculated between $x = 0{\cdot}9$ and $x = 1{\cdot}0$ in the usual manner was 30 dB. The "true reflection coefficient" at the height corresponding to $x = 0{\cdot}9$ is according to equation (6) (in logarithmic form) the sum of these two results or 45 dB.

Fig. 1 gives a useful intuitive picture of the reflection process though it would be wrong to attach too exact an interpretation to the quantity R except in regions

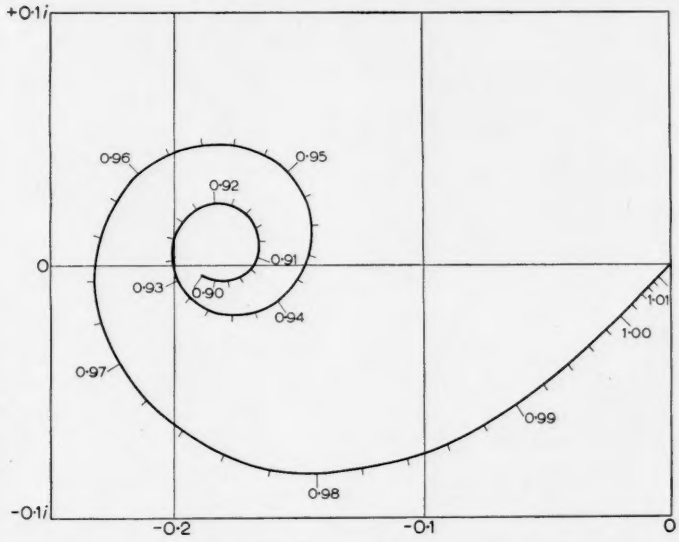


Fig. 1. Complex values assumed by the variable R during the integration process for values of x between 1·01 and 0·90.

where it is constant and therefore the W.K.B. approximation is valid. There is thus a slight inexactitude in the statement about the "true" reflection coefficient at $x = 0{\cdot}9$ since the spiral has not yet come to a point.

3. Calculation of Absorption for an Assumed Model

The electron densities and collision frequencies for the model used in this section are shown by Fig. 2. They are believed to apply approximately to the lower ionosphere over Johannesburg on a summer midday near sunspot maximum. The electron densities below 80–85 km are derived from observations of weak echoes from the D-region published in a companion paper (Fejer and Vice, 1959). The electron densities above 85 km are based on rocket measurements (Jackson and Seddon, 1958, and previous publications) above White Sands, New Mexico, on a summer midday. The collision frequencies are based on laboratory measurement by Crompton *et al.* (1953).

Fig. 3 shows the absorption calculated for this model, integrated downward from the theoretical height of reflection, as a function of the height at which the

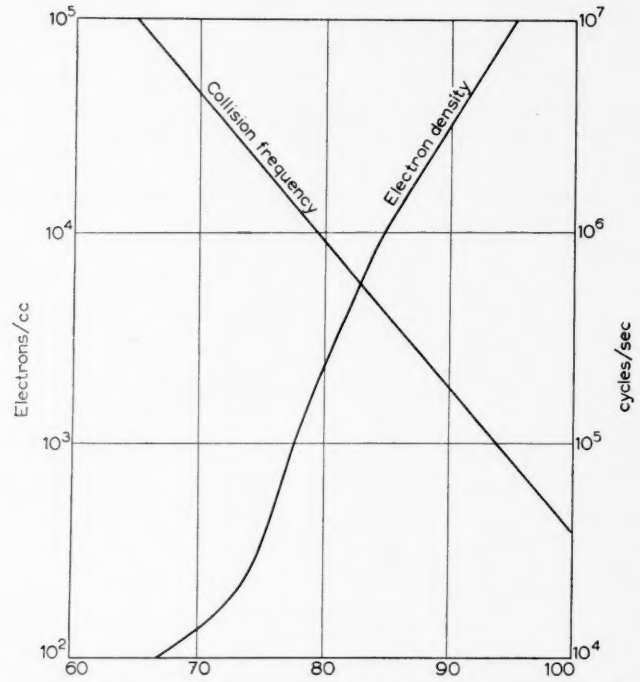


Fig. 2. Assumed variations of electron density and collision frequency with height.

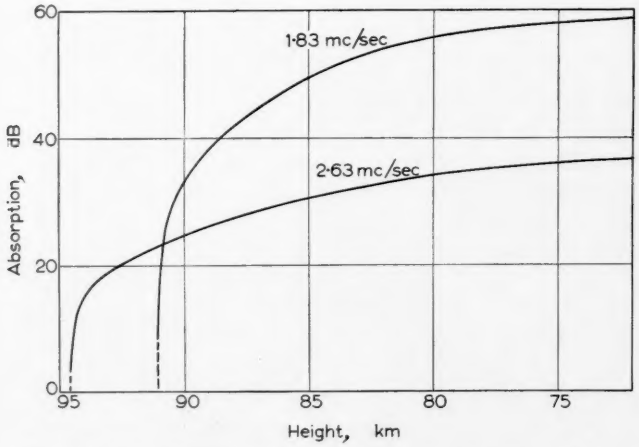


Fig. 3. Absorption integrated downward from the point of reflection as a function of the height.

VOL. 16 1959

integration was stopped. The full-wave correction is shown as a discontinuity at the theoretical height of reflection. The ordinate in Fig. 3 is thus the reflection coefficient defined by equation (6). The total calculated absorption is 58·8 dB for 1·83 Mc/s and 36·7 dB for 2·63 Mc/s. These two frequencies were used in our computations because results of absorption measurements for Johannesburg are available on these two frequencies. The measured monthly median values of absorption at midday for January and February 1958 were 60 dB and 49 dB for 1·83 Mc/s, and 47 dB and 41 dB for 2·63 Mc/s.

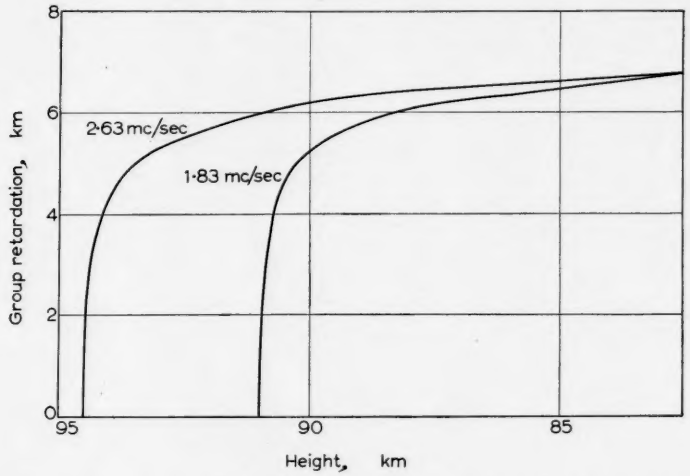


Fig. 4. Group retardation integrated downward from the point of reflection as a function of the height.

As shown in Fig. 3, only 8 dB on 1·83 Mc/s and 3 dB on 2·63 Mc/s of the total absorption is due to the full wave correction. Fig. 3 also show that more than half the absorption occurs within 2 km of the theoretical height of reflection. The slope of the refractive index profile near this height has, therefore, a very great influence on the value of the total absorption.

The discrepancies between the observed values of absorption and those resulting for the present model are not very great, and a relatively small modification in the slope of the electron density profile at the two reflection points would eliminate them.

A further check on the correctness of the model may be carried out by calculating the group delays using the tables of SHINN (1955) for our two frequencies. The errors caused in the calculation by using these tables, which are based on the ray approximation, were found negligible in a typical case by evaluating the group retardation from the difference in phase delay of two full-wave solutions for almost identical frequencies.

The group delay integrated downward from the theoretical height of reflection is shown for two frequencies by Fig. 4. More than three-quarters of this retardation is seen to occur within 2 km of the theoretical height of reflection. The calculated total retardations are 6·8 km for 1·83 Mc/s, and 6·75 km for 2·63 Mc/s. The

resulting virtual heights are 97·8 km for 1·83 Mc/s, and 101·3 km for 2·63 Mc/s. The measured monthly median values of virtual height were 99·5 and 99 km on 1·83 Mc/s, and 103 and 102·5 km on 2·63 Mc/s for the months of January and February 1958, after correcting for the altitude of Johannesburg. The accuracy in the measurement of virtual height was about ± 1 km and the discrepancies are therefore not very significant. A reduction in the slope of the assumed electron density profile near the 2·63 Mc/s reflection point would increase both the absorption and the group retardation and thus reduce the discrepancy in both.

4. DISCUSSION

It was shown in the previous section that measured values of absorption and virtual height are fairly well explained by an ionospheric model based on measurements of an entirely different nature. The question arises whether the measured values of absorption and virtual height for a summer midday could not be equally well explained by some other model, possibly based on data from other workers.

One could for example assume the collision frequency profile proposed by NICOLET (1953). This would mean collision frequencies higher by a factor of about 4, and if the rocket model of electron densities is still assumed, many hundreds of decibels of absorption result. One would have to assume a very much steeper slope of electron density at the reflection heights corresponding to the two frequencies. This in turn would lead to very much smaller group retardations and therefore to greater assumed reflection heights, and would further help to reduce absorption by reducing the collision frequency near the reflection height. There is, however, a difficulty in these assumptions. With smaller group retardations the measured difference between virtual heights would approximately indicate the actual difference between the reflection heights and thus the slope in electron density. There is only one way left out of this dilemma. One could assume that there is a flat section between the two steep sections of the electron density profile. This assumption, which appears rather artificial, as the two frequencies are rather close to each other, would mean that the absorption and the group retardation would have to be extremely high at an intermediate frequency; this is not observed.

Having shown that the collision frequency can not be very much higher it would be appropriate to consider very much lower collision frequencies than those assumed for our model in Section 3. The possibility of a substantial part of the absorption occurring below a height of about 85 km can be discounted in view of the results of investigations described in the companion paper (FEJER and VICE, 1959). In view of this, a less steep slope of electron density would have to be assumed, resulting in much greater group retardations than 7 km. This would put the actual height of reflection for the lower frequency to about 85 km for only twice as great a group retardation as that of the model of Section 3, contradicting the experimental results of the companion paper. A smaller overall slope would also require more delay, and hence smaller slope, on 1·83 Mc/s than on 2·63 Mc/s. This, combined with the greater height difference, and therefore collision frequency difference, would increase the absorption difference which is already too high.

VOL. 16 1959

Further very strong evidence in favour of our model, with relatively large group retardation and low collision frequencies, is provided by the correlation shown in Fig. 5 between the measured differences of absorption and the differences in virtual height at our two frequencies. Pairs of measured absorption differences and virtual height differences near midday are shown by the points of Fig. 5 for the first 5 months of 1958. Each point represents pairs of samples of twenty-four readings each taken during the same quarter of an hour for both frequencies.

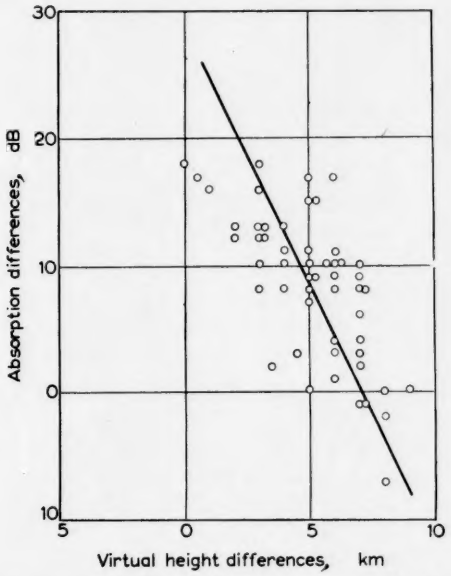


Fig. 5. Correlation between the measured difference in absorption (taking the 1·83 Mc/s value with a positive sign and the 2·63 Mc/s value with a negative sign) and the measured difference in virtual height (taking the 2·63 Mc/s value with a positive sign and the 1·83 Mc/s with a negative sign) on two frequencies.

Only those pairs of samples for which the sum of the measured median values of absorption was within the range of 85–115 dB, were plotted in Fig. 5.

Normally absorption is greater on the lower frequency while the virtual height is greater on the higher frequency, the differences being regarded positive under these conditions. Both a decrease in the slope of the assumed electron density profile at the reflection point of the higher frequency, or an increase in the slope at the reflection point of the lower frequency, lead to a decrease in the absorption difference and to an increase in the virtual height difference. This sense of correlation is shown by Fig. 5.

If the variation in virtual height difference was caused by a variation in the difference in reflection heights rather than group retardations, then a large height difference would be associated with a large difference of collision frequencies and therefore a large absorption difference. Thus if a correlation existed in such a case, it would have the opposite sense to that shown in Fig. 5.

The correlation coefficient calculated from the points of Fig. 5 is 0·70 and

Student's t-test shows it to be very highly significant. The straight line in Fig. 5 is a regression line of the absorption difference on the height difference.

Summing up, it appears that the present model follows approximately from the absorption and virtual height measurements, and the results of the companion paper. It is hoped to extend this model, at least tentatively, to different times of the day and the year on the basis of the results of absorption measurements, which will be published in the near future. It is regrettable that absorption and virtual height measurements were restricted to two frequencies. Measurements on a larger number of frequencies would be very desirable with the aim of determining the shape of the lower E-region.

One of the difficulties in any programme of this nature would be the interpretation, on account of the mathematical complexity of the problem. It takes a few days to determine the absorption for a given electron density and collision frequency profile using a desk calculator. Most of this time is taken up by the determination of n and NN''; the actual numerical integration of the differential equation only needs a few hours.

The computed full wave corrections to the absorption are relatively small though not negligible. The correction may well be larger nearer to the critical frequency, at lower frequencies or with a different inclination of the magnetic field of the earth, which has a large influence on the absorption near the reflection point.

VOL. 16 1959

5. CONCLUSIONS

The relatively simple method of obtaining full wave solutions described in this paper increases the accuracy of absorption calculations. The time required to calculate the integrated absorption with the help of the Appleton–Hartree formula is already considerable and the additional time required to calculate the full wave correction is not very great.

The application of the method is not restricted to the calculation of ionospheric absorption. Solutions of the wave equation may be required for an entirely different purpose.

The method has two distinct advantages compared with a similar method given by SCHELKUNOFF; it is less susceptible to accumulated errors and its application may be restricted to a narrower range.

Measurements of monthly median values of absorption and virtual height in Johannesburg on summer middays are reasonably well explained by a model of the lower ionosphere based on rocket measurement and certain investigations described in a companion paper (FEJER and VICE, 1959).

The desirability of absorption and virtual height measurements on more than two frequencies, possibly combined with the type of investigations described in the companion paper, for the purpose of determining electron densities and collision frequencies in the lower ionosphere, is stressed.

Acknowledgements—The authors would like to thank the South African Council for Scientific and Industrial Research for permission to publish this paper. This work formed part of the South African programme for the International Geophysical Year.

References

Appleton E. V.	1932	*J. Instn. Elect. Engrs.* **71,** 642.
Beynon W. J. G. and Davies K.	1955	*The Physics of the Ionosphere* p. 40. Physical Society, London.
Crompton R. W., Huxley L. G. H. and Sutton D. J.	1953	*Proc. Roy. Soc.* A**218,** 507.
Fejer J. A.	1955	*J. Atmosph. Terr. Phys.* **7,** 322.
Fejer J. A. and Vice R. W.	1959	*J. Atmosph. Terr. Phys.* **16,** 291.
Försterling K.	1942	*Hochfrequenztech. u. Elektroakust.* **59,** 10.
Gardner F. F. and Pawsey J. L.	1953	*J. Atmosph. Terr. Phys.* **3,** 321.
Gibbons J. J.	1955	*The Physics of the Ionosphere* p. 351. Physical Society, London.
Jackson E. J. and Seddon J. C.	1958	*J. Geophys. Res.* **63,** 197.
Jeffreys H.	1924	*Proc. Lond. Math. Soc.* (2) **23,** 428.
Millington G.	1938	*Proc. Phys. Soc. Lond.* **50,** 561.
Mitra S. K.	1952	*The Upper Atmosphere* (2nd Ed.) p. 245. The Asiatic Society, London.
Nicolet M.	1953	*J. Atmosph. Terr. Phys.* **3,** 200.
Schelkunoff S. A.	1951	*Theory of Electromagnetic Waves* Interscience New York.
Shinn D. H.	1955	*The Physics of the Ionosphere* p. 402. Physical Society, London.

DL.
16
959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 318 to 323. Pergamon Press Ltd. Printed in Northern Ireland

On a pseudo-iterating method for reduction of observations of earth-tide for effects from tides of the seas, with an application to observations made at Bergen (Norway) 1934

J. EGEDAL
Kongevejen 126, Birkerød, Denmark

(*Received* 22 *January* 1959)

Abstract—A new method for the reduction of observations of earth-tide for effects from tides of the sea is given for a special case. The method is applicable only when the variation of the effect from the tides of the sea is different from the variation of the earth-tide. In cases where the variation of the effect from the tides of the sea is difficult to determine with certainty, or where the observations themselves are uncertain, the method may be inapplicable.

1. INTRODUCTION

IF observations of tidal movements of the earth's crust are made at a place where a perceptible effect from the tides of the sea exists, it is necessary to reduce the observations for this effect in order to get the real effect of the sun and the moon on the earth's crust. This reduction should also be made before the determination of *the diminution factor*, D, namely the factor by which the theoretical effect on an unyielding earth should be multiplied in order to give the resulting effect on the yielding earth. For this reduction of tidal movements of the earth's crust a *pseudo-iterating method* may in some cases be used. Below, the method is described and, as an example, used for the elimination of tidal effects from the sea on observations made at the Geophysical Institute at Bergen.

VOL. 16 1959

2. THE METHOD

The observed variation of the movements of the earth's crust, T, consists of variations caused by the direct lunar and solar effects and by the tides of the sea, T_s. The direct effect can be expressed by:

$$DT_d$$

where D is the diminution factor and T_d the theoretical effect on an unyielding earth, and the following equation can be written down:

$$DT_d + T_s - T = 0. \tag{1}$$

In equation (1), D is the unknown to be determined, T_d a quantity which can be computed, T_s an unknown quantity giving the effect of the tides of the sea at the place of observations, and T the resulting effect.

In cases where T_d and T_s vary in a rather different way the pseudo-iterating method may be used.

From determinations of later years the value of D seems to be greater than 0·70 and less than 0·75. A value which is, say 0·70, may be substituted for D in equation (1) for the determination of T_s for all full hours of the day to be examined. By

this procedure an error is introduced in the determination of T_s, but if T_d for example, contains diurnal as well as semi-diurnal terms, and T_s only semi-diurnal terms, an adjustment of the found hourly values of T_s will eliminate the contribution to the error from the diurnal term of T_d, so that only a contribution from the semi-diurnal term affects T_s. The variations are supposed to take place during a lunar day. By means of the adjusted hourly values for T_s a new value for D is determined. In certain cases this new value of D will be nearer to the correct value for the given observations than the applied value. The conditions on which this takes place will now be elucidated.

Let D be the value which would be found from a correct treatment of the given observations, and ΔD a quantity amounting to some few per cent of D. The value $D' = D - \Delta D$, where ΔD is supposed to be a positive quantity, is substituted for D in (1), and the value of T_s is determined. The twenty-four quantities T_d are supposed to be sums of contributions from diurnal and semi-diurnal terms:

$$T_d = m_1 \cos(rt + \alpha_1) + m_2 \cos(2rt + \alpha_2); \quad r = 0, 1, 2, \ldots, 23 \tag{2}$$

where r is the hour, and $t = \frac{2\pi}{24}$.

The quantities T_s is supposed to be represented by a semi-diurnal term:

$$T_s = m_2' \cos(2rt + \beta_2); \quad r = 0, 1, 2, \ldots, 23 \tag{3}$$

where r is the hour.

The equations (1), (2) and (3) give:

$$D[m_1 \cos(rt + \alpha_1) + m_2 \cos(2rt + \alpha_2)] + m_2' \cos(2rt + \beta_2) - T_r = 0$$
$$r = 0, 1, 2, \ldots, 23. \tag{4}$$

Equation (4) comprises equations for each of the 24 hr of the considered day.

In order to get an approximate value for the term:

$$m_2' \cos(2rt + \beta_2)$$

D in (4) is substituted by the too low value $D' = D - \Delta D$:

$$D[m_1 \cos(rt + \alpha_1) + m_2 \cos(2rt + \alpha_2)] - \Delta D[m_1 \cos(rt + \alpha_1) + m_2 \cos(2rt + \alpha_2)]$$
$$+ m_2'' \cos(2rt + \beta_2') - T_r = 0; \quad r = 0, 1, 2, \ldots, 23 \tag{5}$$

where m_2'' and β_2' are the determined values.

From (4) and (5) is obtained

$$m_2'' \cos(2rt + \beta_2') = m_2' \cos(2rt + \beta_2) + \Delta D[m_1 \cos(rt + \alpha_1) + m_2 \cos(2rt + \alpha_2)]$$
$$r = 0, 1, 2, \ldots, 23. \tag{6}$$

First the contribution of the term $\Delta D m_1 \cos(rt + \alpha_1)$ shall be considered. As it is a diurnal term the contribution to the semi-diurnal term will be negligible. From

$$m_2'' \cos(2rt + \beta_2') = m_2' \cos(2rt + \beta_2) + \Delta D m_2 \cos(2rt + \alpha_2) \tag{7}$$

is obtained:

$$m_2'' = m_2' + \Delta D m_2 \cos(\alpha_2 - \beta_2), \tag{8a}$$

when small quantities are neglected, and

$$\operatorname{tg} \beta_2' = \frac{m_2' \sin \beta_2 + \Delta D m_2 \sin \alpha_2}{m_2' \cos \beta_2 + \Delta D m_2 \cos \alpha_2}. \tag{8b}$$

Hereafter (5) may be written in the following form:

$$D[m_1 \cos(rt + \alpha_1) + m_2 \cos(2rt + \alpha_2)] + [m_2' + \Delta D m_2 \cos(\alpha_2 - \beta_2)] \times \cos(2rt + \beta_2') - T_r = 0; \quad r = 0, 1, 2, \ldots, 23. \tag{9}$$

By means of the least square method a new value D'' will be determined from (9) as a function of ΔD.

The form of the normal equation is as follows:

$$\sum_0^{23} [m_1 \cos(rt + \alpha_1) + m_2 \cos(2rt + \alpha_2)] \times [m_1 \cos(rt + \alpha_1) + m_2 \cos(2rt + \alpha_2)] D'' + \sum_0^{23} [m_1 \cos(rt + \alpha_1) + m_2 \cos(2rt + \alpha_2)] \times [m_2' + \Delta D m_2 \cos(\alpha_2 - \beta_2) \cos(2rt + \beta_2') - T_r] = 0 \tag{10}$$

or

$$[\sum_0^{23} m_1^2 \cos^2(rt + \alpha_1) + 2\sum_0^{23} m_1 m_2 \cos(rt + \alpha_1) \cos(2rt + \alpha_2) + \sum_0^{23} m_2^2 \cos^2(2rt + \alpha_2)] D'' + \sum_0^{23} m_1 m_2 \cos(rt + \alpha_1) \cos(2rt + \beta_2') + \sum_0^{23} m_1 \Delta D m_2 \cos(rt + \alpha_1) \cos(\alpha_2 - \beta_2) \cos(2rt + \beta_2') - \sum_0^{23} m_1 T_r \cos(rt + \alpha_1) + \sum_0^{23} m_2 m_2' \cos(2rt + \alpha_2) \cos(2rt + \beta_2') + \sum_0^{23} m_2^2 \Delta D \cos(\alpha_2 - \beta_2) \cos(2rt + \alpha_2) \times \cos(2rt + \beta_2') - \sum_0^{23} m_2 T_r \cos(2rt + \alpha_2) = 0. \tag{11}$$

Neglecting all small terms the following equation is obtained:

$$[12m_1^2 + 12m_2^2] D'' - \sum_0^{23} m_1 T_r \cos(rt + \alpha_1) + 12 m_2 m_2' \cos(\alpha_2 - \beta_2') + 12 m_2^2 \Delta D \cos(\alpha_2 - \beta_2) \cos(\alpha_2 - \beta_2') - \sum_0^{23} m_2 T_r \cos(2rt + \alpha_2) = 0 \tag{12}$$

and

$$D'' = -\frac{1}{12m_1^2 + 12m_2^2} [-\sum_0^{23} m_1 T_r \cos(rt + \alpha_1) + 12 m_2 m_2' \cos(\alpha_2 - \beta_2') + 12 m_2 \Delta D \cos(\alpha_2 - \beta_2) \cos(\alpha_2 - \beta_2') - \sum_0^{23} m_2 T_r \cos(2rt + \alpha_2)]. \tag{13}$$

VOL. 16 1959

Putting $\Delta D = 0$ the value D can be found:

$$D = -\frac{1}{12{m_1}^2 + 12{m_2}^2}[-\sum_0^{23} m_1 T_r \cos(rt + \alpha_1) + 12 m_2 m_2' \cos(\alpha_2 - \beta_2) - \sum_0^{23} m_2 T_r \cos(2rt + \alpha_2)]. \quad (14)$$

From (13) and (14) the following result is obtained:

$$D'' = D - \Delta D \frac{{m_2}^2 \cos(\alpha_2 - \beta_2) \cos(\alpha_2 - \beta_2')}{{m_1}^2 + {m_2}^2}. \quad (15)$$

As β_2 and β_2' differ only a little (cf. (8b)) the factor to ΔD, except for values of $\alpha_2 - \beta_2$ and $\alpha_2 - \beta_2'$ near to 90° and 270°, will be positive and below 1, and under these circumstances D'' is greater than D' which is equal to $D - \Delta D$.

By the example it is shown that in most cases the procedure results in a value for D nearer to the correct value for the given observations than that with which the computations are started.

Continuing the procedure starting with D'' the correct value of D for the given observations may gradually be determined, but instead of continuing the iteration after having made the computations twice it is easier to estimate the final value and test its correctness.

If the computation has been started with a too high value of D, corresponding to a negative value of ΔD, the new value for D would also have been nearer to the correct D for the given observations (cf. (15)).

3. The Observations

The observations to be treated were made at the Geophysical Institute, Bergen ($\varphi = 60°23'\text{N}$; $\lambda = 5°20'\text{E}$). During the period June 10–24, 1934, two *Niveauvariometers* (Egedal, 1947) placed in a tunnel 18m below the surface of the earth at the ends of a 103 m long pipe were read every full hour, day and night (Egedal and Fjeldstad, 1937). The pipe was placed horizontally and half-filled with distilled water. The differences between corresponding readings of the two Niveauvariometers contain the variation of the height difference between the pillars for the Niveauvariometers as well as the difference between the varying zero-points of the Niveauvariometers. An adjusted value of the difference of the zero-points has been determined as the mean value of the obtained differences of an interval of 24hr 50min centred at the considered hour. The differences are given in tenths of μ ($1\ \mu = 0{\cdot}001$ mm).

4. Application of the Pseudo-iterating Method to the Observations

The data to which the pseudo-iterating method should be applied consist of hourly values for the whole period of observations from 10 June to 24 June.

As starting value for D, 0·700 is chosen, the variation of the differences for an unyielding earth are computed, and thereafter the hourly values of T_s are determined by means of (1) for each of the 15 days. For each day an amplitude and a phase-angle of T_s are found for a lunar semi-diurnal term. The amplitudes are adjusted, and the phase-angles are corrected in order to vary as the tide of the sea, so that adjusted hourly values of T_s for the whole period are at hand. These hourly

values are subtracted from the corresponding hourly values derived from the observations, and from the resulting hourly values the quantity D'' is finally determined for each day by means of the least square method:

starting value for D	resulting value for D
0·700	0·717

Now the procedure is continued using as starting value for D: 0·717, and the following result is obtained:

starting value for D	resulting value for D
0·717	0·726

Instead of continuing the iteration till the starting and the resulting values are equal it is preferable to estimate a final value and to test its correctness. In the present case $D = 0{\cdot}733$ was chosen.

In Table 1 the found values of D for each of the days and their standard deviations are given.

Table 1. Found values of D and their standard deviations

Date	D	Date	D
1934 June 10	0·712 ± 0·043	June 18	0·428 ± 0·106
11	0·782 ± 0·031	19	0·692 ± 0·069
12	0·760 ± 0·025	20	0·590 ± 0·058
13	0·747 ± 0·027	21	0·721 ± 0·048
14	0·768 ± 0·022	22	0·697 ± 0·039
15	0·660 ± 0·031	23	0·701 ± 0·055
16	0·734 ± 0·043	24	0·658 ± 0·046
17	0·875 ± 0·054		

The weighted mean of the 15 values of D in Table 1 is

$$D = 0{\cdot}733 \pm 0{\cdot}016.$$

This value is the final value, as it is equal to the starting value.

In order to elucidate the quality of the observations and the magnitude of the variations of T_s and DT_d, Fig. 1 and Fig. 2 are given. Both figures give the variations for the same day, i.e. 12 June 1934.

In Fig. 1 the circles represent hourly values equal to the derived hourly values minus $0{\cdot}733T_d$, and the curve gives the semi-diurnal variation for the considered day determined on the basis of the above mentioned adjusted amplitudes and corrected phase angles for T_s.

In Fig. 2 the circles represent hourly values equal to the derived hourly values minus corresponding values given by the curve of Fig. 1. The curve of Fig. 2 represents $0{\cdot}733T_d$.

The values represented by the circles in the two figures are found by means of equation (1), and therefore their departures from the curve are equal for the same hour in Fig. 1 and Fig. 2.

VOL. 16 1959

5. Conclusive Remarks

Usually the treatment of observations of earth-tide consists in an examination of some of the more important tidal constituents. In the present case the derived variation is treated as a whole, and in this way the quantities to be examined are greater than in the case where tidal constituents are to be examined. This is an advantage and a simplification.

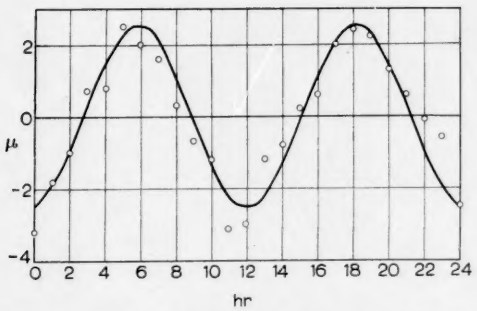


Fig. 1. The effect of the tides of the sea on 12 June, 1934. The circles give hourly values computed from the observations, and the curve is derived on the basis of observations during 15 days.

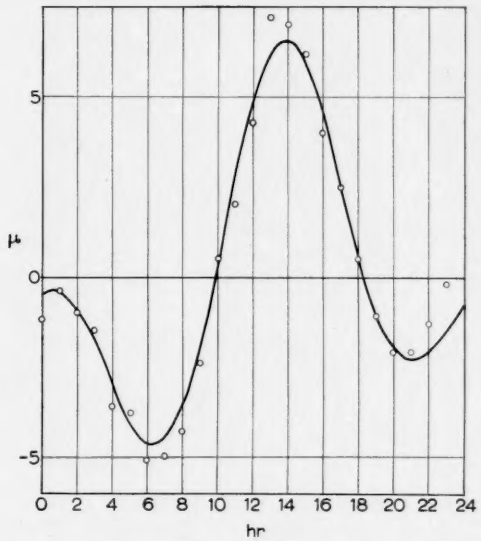


Fig. 2. The height differences caused by earth-tide on 12 June, 1934. The circles give the hourly values after the elimination of the effect from the tides of the sea, and the curve gives the effect on an unyielding earth multiplied by the diminution factor (0·733).

In Section 2 the method is described and an example based on certain assumptions and simplifications is given. These assumptions are chosen, so that they are as far as possible in accordance with the circumstances at Bergen. Therefore, the example demonstrates that the iteration is applicable only for the case examined. In other places of observations it must first be examined whether or not the method is applicable.

The variation of the effect from the tides of the sea may be difficult to estimate. The area from which appreciable effects arise is large and the variation of the tides within the area may therefore be different. If so, the determination of the variation of the effect from the tides of the sea may be uncertain. In the case dealt with here the variation was rather simple as the tides at the Norwegian coast are mainly semi-diurnal.

References

Egedal J. 1947 *Publ. Danske Met. Inst.* No. 10.
Egedal J. and Fjeldstad J. E. 1937 *Geofys. Publ.* **11**, No. 14.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 324 to 339. Pergamon Press Ltd. Printed in Northern Ireland

Die Frequenzabhängigkeit der ionosphärischen Absorption*

Julius Bartels zum 60. Geburtstag Gewidmet

K. Bibl, A. Paul und K. Rawer
Ionosphären-Institut Breisach, im Fernmeldetechnischen Zentralamt der Deutschen Bundespost

(*Received* 5 *February* 1959)

Abstract—Detailed analysis has shown that in most cases where observations did not follow the writers' old formula $L = B/(f + f_L)^2 + N \cdot \Delta_3(f_0E/f)$, intense echos from a blanketing E_s-layer were present. The theory is generalized to the case of a thin, mirror-like E_s-layer embedded into a parabolic E-layer (model **D**). This gives a new function depending on a second parameter given by the altitude of the mirror. By application of the new formula (with the actual parameters of each observation) we obtain a much better description of the observed frequency variation. Noon values of B (non-deviative D-absorption) and N (deviative E-absorption) show opposed seasonal variations, N increasing but B decreasing with the sun's height. The variation of N can be explained satisfactorily by a small decrease of the altitude of the E-layer with increasing height of the sun.

1. Bisherige Theorien

Die Impulsmessungen der ionosphärischen Absorption auf fester Frequenz nach der Methode des Reflexionskoeffizienten (Ratcliffe, 1935; Piggott, 1953; Rawer, 1951) werden neuerdings im Rahmen des AGI auf einer Reihe von Stationen regelmässig mittags ausgeführt. Dagegen wird die Frequenzabhängigkeit der Absorption nur an wenigen Stellen beobachtet, weil einerseits diese Messungen grösseren technischen Aufwand erfordern, andererseits der Frequenzgang bisher noch keine voll befriedigende Erklärung gefunden hatte. Unter diesen Umständen sieht man im AGI davon ab (Piggott *et al.*, 1957), die früher übliche Reduktion der Messwerte auf eine Normalfrequenz weiterhin durchzuführen.

Lange Zeit wurden die Ergebnisse der Absorptionsmessungen mit einer einfachen Formel interpretiert, die lediglich eine brechungsfreie (non-deviative) Absorption in der D-Region annimmt (Modell **A**). Da dort voraussetzungsgemäss die Elektronendichte gering ist, kann die quasilongitudinale Approximation benutzt werden, meist in Bookers (1935) Form

$$L = A/(f \pm f_L)^2 \tag{1}$$

L ist das Absorptionsdekrement (hier immer dB), f die Frequenz, f_L die longitudinale Komponente der Gyrofrequenz f_H. Das obere Vorzeichen entspricht der ordentlichen Komponente, für die die Messungen im allgemeinen ausgeführt werden.

Im Anhang wird ausgeführt, dass eine konsequentere QL Näherung lauten sollte $A/(f \pm f_H)^2$. Der Unterschied beider Formeln fällt in mittleren Breiten nicht sehr ins Gewicht, ist aber für niedere Breiten nicht unbedeutend. Zwei von uns (Bibl und Rawer, 1951) haben später ein vollständigeres Modell vorgeschlagen (Modell **B**), in dem neben der nicht-selektiven Absorption in der D-Region auch

* Die vorliegende Untersuchung wurde durch eine Forschungsbeihilfe des Bundesministeriums für Wirtschaft ermöglicht Kontrakt Nr. A 191.

VOL. 16 1959

die selektive (deviative) Absorption in der normalen E-Schicht berücksichtigt wird. Sie benützten für letzetere ein parabolisches Modell der E-Schicht mit exponentieller Abnahme der Stosszahl und schrieben für das Absorptionsdekrement der ordentlichen Komponente:

$$L = B/(f + f_L)^2 + N \,.\, \Delta_3(\eta) \tag{2}$$

wobei $\eta = f_0E/f$ das Reziproke der auf die kritische Frequenz der E-Schicht, f_0E, bezogenen Frequenz ist. Die Funktion im zweiten Term, der die selektive Absorption beschreibt, ist das Integral über den Absorptionskoeffizienten längs des Lotungsweges in der brechenden E-Region. Das Integral wurde einerseits durch Reihenentwicklung numerisch berechnet (Bibl und Rawer, 1951), andererseits konnte es mit Hilfe spezieller transzendenter Funktionen angegeben werden (Argence *et al.*, 1950). Dabei wurde im zweiten Term die Sellmeier-sche Dispersionsformel benutzt, also der Einfluss des magnetischen Feldes der Erde vernachlässigt.

Ein dritter Vorschlag, weniger auf einem physikalischen Modell, als auf einer numerischen Analyse von Messdaten beruhend, wurde von McAllcock (1955) gemacht. Er schreibt, in erster Linie für den Bereich sehr niedriger Frequenzen $L = c_1 \,.\, \log f + c_2$. Wir werden diesen Vorschlag hier nicht näher diskutieren, da uns die Messungen im Bereich oberhalb 1,5 MHz interessieren werden.

Ein Vorschlag von Beynon und Davies (1955) führt die selektive Absorption in einer Chapman-Schicht, durch eine Funktion F $(1/\eta)$ gegeben, ein und kombiniert damit eine frequenzunabhängige Konstante:

$$L = k_1 + k_2 \,.\, F(1/\eta) \,.\, \cos\chi$$

χ ist die Zenithdistanz der Sonne, die Funktion F liegt tabuliert vor; als einfache, nicht-selektive Näherung für F wird f^{-2} angegeben. Die Verfasser wollen ihre Formel mit dem etwas verblüffenden konstanten Term als eine Näherung aufgefasst wissen für ein der Wirklichkeit besser entsprechendes Schichtmodell.

Eine ähnliche Formel hat Whitehead (1957) benutzt, um Beobachtungen auf zwei Frequenzen darzustellen (Modell **C**):

$$L = B_1/(f + f_L)^2 + C_1 \tag{3}$$

Nach seiner Interpretation wird die Tag-zu-Tag Variation ebenso wie die Winteranomalie von B_1 verursacht.

2. Vergleich mit Messergebnissen

Die seinerzeit für die Jahre 1949–1950 durchgeführten Untersuchungen hatten eine recht erhebliche Streuung von Tag zu Tag gezeigt, so dass darauf verzichtet wurde, den Frequenzgang einer Einzelmessung zu interpretieren. Die Streuung war bei den Messergebnissen von Slough und Freiburg ähnlich breit gewesen, obwohl beide Stationen von einander abweichende Messverfahren benutzen. Die Analyse hatte sich deshalb auf monatliche Medianwerte je Frequenz beschränkt. Es war vermutet worden, dass die Streuung der Einzelwerte im wesentlichen von Fokussierung herkomme.

Die von Harnischmacher und Rawer entwickelte Ionosonde kann direkt zur Absorptionsmessung benutzt werden, weil ihr Empfänger eine logarithmische

Empfindlichkeitskurve hat und seine Ausgangsspannung eine eindeutige Funktion der hochfrequenten Eingangsspannung ist. Da man für Absorptionsmessungen längere Zeit auf derselben Frequenz beobachtet, um Fadings auszumitteln und um definierte Abstrahlungsverhältnisse zu haben, war schon in der ursprünglichen Ausführung dieser Ionosonde die mechanische Möglichkeit vorgesehen, die Bewegung des Drehkondensators mit Hilfe einer magnetischen Kupplung anzuhalten. Dann wird die Frequenz nur noch unstetig durch den Spulenwechsel bei der Bereichschaltung verändert. Entsprechend den 8 Spulensätzen des Gerätes erhält man 8 feste Frequenzen; früher wurde jede Frequenz 75 sec lang betrieben, so dass ein voller Zyklus 10 min dauerte, neuerdings haben wir auf 3 min-Zyklus umgestellt, so dass jede Frequenz 22,5 sec steht. Kürzlich hat BIBL mit Relais zusätzliche Abstimm-Mittel jeweils eingeschaltet, so dass die Wahl der Messfrequenzen besser an die Bedürfnisse angepasst werden konnte. Die Messfrequenzen sind seit 1957: 1,72; 2,05; 2,44; 2,90; 3,45; 4,10; 4,88; 5,80 MHz und liegen enger gestaffelt als früher. In früherer Zeit haben wir die Amplituden der verschiedenen Echos alle 10 sec von einem Beobachter ablesen lassen; jede Frequenz wurde zwei mal im Abstand von 10 min 75 sec lang beobachtet. Wir erhalten neuerdings durch die viermalige Wiederholung der Beobachtung innerhalb 12 min eine statistisch zuverlässigere Auskunft über die Amplituden. Während frühere Versuche, die Ausmittelung der Amplituden durch fotographisches Überschreiben* zu erhalten, fehlgeschlagen sind, gelang es BIBL jetzt mit 22,5 sec Überschreibungszeit bei höchster Konstanz der Betriebsbedingungen (Verstärkung sowie Schärfe und Helligkeit der Registrierröhre) Registrierungen zu erhalten, deren Auswertung zufriedenstellend mit dem Ergebnis von Amplitudenablesungen übereinstimmt.

VOL. 16 1959

Zur Auswertung wird das Stehbild projiziert und im allgemeinen aus Schwärzungs-Schwerpunkt und geometrischem Mittel der Extremalwerte eine mittlere Amplitude bestimmt. In 90% der Fälle stimmt dieser Amplitudenwert mit dem zeitlichen Medianwert genügend genau überin. In den wenigen Zweifelsfällen, etwa wenn zwei Schwärzungs-Schwerpunkte auftreten, wird versucht, aus der Schwärzungsverteilung auf die Lage des Medianwertes zu schliessen. Die Registriermethode erfordert zwar sorgfältige Wartung, hat aber die Vorteile, dass die Beobachtungen automatisch anfallen und die Unterlagen jederzeit nachgeprüft werden können.

Die Beobachtungen im geophysikalischen Jahr wurden mit der neuen Registriermethode durchgeführt und für die Eichung die Konventionen (PIGGOTT *et al.*, 1957) befolgt; demgemäss wird durch Vergleich mit ungestörten Nachtwerten geeicht, während wir dies früher durch Vergleich mit Mehrfachechos getan hatten.

Inzwischen haben wir eine grosse Zahl von Mittagsbeobachtungen einzeln untersucht und mit den zugehörigen Ionogrammen verglichen. Dabei ergab sich, dass gewisse Typen der sporadischen E-Schicht, E_s, erhebliche Abweichung vom normalen Frequenzverlauf der Echo-Amplitude bewirken. Die selektive Absorption in der E-Schicht kann nicht geleugnet werden, aber sie wird nur in den Fällen klar beobachtet, in denen keine E_s-Schicht vorhandcen ist.

* Photographic integration

Schon bei der bisherigen Auswertung der Messungen waren die Fälle ausgeschlossen worden, in denen eine halbdurchlässige E_s-Schicht auftrat. Dagegen waren Fälle gut reflektierender E_s-Schichten nicht besonders unterschieden worden; es ist tatsächlich unmöglich, aus einer Einzelbeobachtung auf fester Frequenz festzustellen, ob ein Echo von der E-oder der E_s-Schicht kommt. Unsere Untersuchung ergab nun, dass im letzten Fall das Echo immer kräftiger war, als wenn die gleiche Frequenz von der normalen E-Schicht reflektiert wurde. Bei Anwesenheit einer derartigen E_s-Schicht ist also für die von ihr reflektierten Frequenzen die Absorption geringer als im Normalfall. Das bedeutet eine einseitige Abweichung vom Normalfall, dem unser Modell **B** entsprach und hat zur Folge, dass die monatlichen Medianwerte nicht mehr einem klaren Modell entsprechen.

Als Beispiel für dieses Verhalten geben wir in Fig. 1 eine Amplitudenregistrierung mit stetig veränderter Frequenz wieder.*

Bei dieser Registrierung wird die Amplitude (in logarithmischem Maßstab) von jedem Echozeichen aus nach unten aufgetragen. Man erkennt in Fig. 1 deutlich die beiden Komponenten des E_s-Echos (top Frequenzen f_0E_s und f_xE_s) und die Abdeckung der F-Region. Diese Abdeckung ist vollständig, so dass $f_bE_s = f_0E_s$; die E_s-Schicht ist also nicht transparent. Der Amplitudenverlauf längs der E-Spur zeigt, beginnend mit niedrigen Frequenzen, zunächst eine Zunahme; sie entspricht dem Frequenzgang der nicht-deviativen Absorption in tiefen Schichten. So ist dann die Frequenz maximaler Verzögerung (cusp frequency) zu erkennen, die in Fig. 1 mit f_r bezeichnet ist. Bei dieser Frequenz geht das Echo von der E-Schicht zur E_s-Schicht über. Von hier ab steigt die Amplitude rasch auf einen grossen Wert an und behält ihn praktisch bis zum Ende des Echozuges bei f_0E_s bei. Der gleiche Verlauf wiederholt sich für die x-Spur mit dem Unterschied, dass dort die nicht-deviative Absorption bis zu wesentlich höheren Frequenzen reicht.

3. Ein neues Modell

Der gute Reflexionskoeffizient dieser E_s-Schichten erfordert eine Erklärung und damit eine Abänderung unseres Modelles **B**. In diesen Fällen kommt die selektive (deviative) Absorption der E-Schicht kaum zur Auswirkung, weil der obere Teil des Lotungsweges durch Reflexion an der E_s-Schicht abgeschnitten wird. Dementsprechend ändern wir Modell **B** in ein neues Modell **D** um, indem wir eine spiegelnde E_s-Schicht in die E-Region einbauen. Ein idealer Spiegel entspricht einer E_s-Schicht mit unstetigem Ionisationsanstieg an der Unterseite. Wir setzen weiter voraus, dass diese E_s-Schicht so dünn ist, dass die vermehrte Ionisation auch für durchgehende Wellen keinen wesentlichen Beitrag zum Absorptionsintegral liefert, dass aber andererseits die Dicke noch gross genug ist, um auf Frequenzen unter der kritischen praktisch totale Reflexion zu erhalten.† Im Modell bedeutet das, dass wir in Höhe der E_s-Schicht einen verlustlosen Spiegel ansetzen, der nur für die Frequenzen zwischen f_r und f_0E_s wirksam ist.

Die Höhe der E_s-Schicht geht in kritischer Weise in unser neues Modell ein. Wenn wir sie aus scheinbaren Höhenablesungen entnehmen würden, wären schon

* Über diese Art der Registrierung wird *A. Paul* demnächst getrennt berichten.

† Homogene E_s-Schichten wie in Fig. 1 sind durchweg so dick, dass nur in unmittelbarer Nähe der kritischen Frequenz partielle Reflexion festgestellt werden kann. Dieser Einfluss bewirkt in Fig. 1 den steilen, aber noch stetigen, Intensitätsabfall bei f_0E_s.

kleine Irrtümer, auch im Verhältnis zur Höhe der E-Schicht, untragbar. Wesentlich sicherer ist es, die relative Höhe aus der Lage der Frequenz f_r (retardation frequency) zu entnehmen. Diese Frequenz bezeichnet ja die Plasmafrequenz in der Höhe der normalen E-Schicht, wo die Unterseite der dünnen E_s-Schicht liegt. Wir haben damit in unserem Modell **D** nun zwei Parameter, nämlich f_0E und f_r. Beide Parameter müssen aus zeitlich benachbarten Ionogrammen entnommen werden, besonders f_r. Die Anwendung der neuen Methode setzt also den Vergleich mit Ionogrammen voraus.

An Stelle einer Ein-Parameter-Funktion $\Delta_3(\eta)$ mit $\eta = f_0E/f$ tritt jetzt eine Zwei-Parameter-Funktion $\Delta_3(\eta;\eta_0)$ mit $\eta_0 = f_0E/f_r$. Die neue Funktion entspricht dem bisherigen Absorptionsintegral mit veränderter oberer Integrationsgrenze. In der Schreibweise von (Bibl and Rawer, 1951) hatten wir als Absorptions-Integral:

$$\Delta_3 = \eta^2 \int_{-1}^{-\sqrt{(1-\eta^{-2})}} \mathrm{d}y \, \frac{(1-y^2)\cdot\psi(y)}{\sqrt{\{-\eta^2(1-y^2)\}}} \tag{4}$$

Wobei die reduzierte Höhe y durch Vergleich mit der Halbdicke der Parabelschicht, Y_m erhalten wird, also $y = Y/Y_m$. ψ entspricht einer reduzierten Stosszahl, die lokale Stosszahl ist $\eta = \eta_0 \cdot \psi(y)$.

Solange die Frequenz so niedrig ist, dass die in der oberen Integrationsgrenze angegebene Reflexionsbedingung erfüllt wird (für den Bereich unterhalb der spiegelnden E_s-Schicht), gilt die bisherige Integrationsgrenze. Liegt die Frequenz jedoch höher, nämlich oberhalb von f_r, so ist als obere Integrationsgrenze $-\sqrt{(1-\eta_0{}^{-2})}$ einzusetzen.

Je nach dem Wert von f_r und damit η_0 ergibt sich eine andere Fortsetzung. Bei der Verzögerungsspitze, die durch f_r (retardation frequency) gegeben ist, wird der regelmässige Verlauf der Absorptionsfunktion abgebrochen.

Fig. 2 gibt die nach der Methode von (Bibl and Rawer 1951) berechneten Werte der Funktion $\Delta_3(\eta;\eta_0)$ wieder, die Begrenzungskurve links oben entspricht der von E_s ungestörten Funktion $\Delta_3(\eta)$.

Wir schreiben dann noch in unserem Modell **D** für das Absorptionsdekrement*

$$L = B/(f+f_H)^2 + N \cdot \Delta_3(\eta;\eta_0) \tag{5}$$

Bei unserer Berechnung wurde wieder die Sellmeier 'sche Dispersionsformel benutzt, wir möchten deshalb noch den eventuellen Einfluss des Erdmagnetfeldes diskutieren.

Vergleicht man die lokalen Werte des Absorptionskoeffizienten mit und ohne Erdmagnetfeld, so findet man recht erhebliche Unterschiede. Es zeigt sich jedoch, dass das Vorzeichen der Differenz verschieden ist in den beiden Booker 'schen Approximationsbereichen QL und QT. Soweit die QL Näherung gültig ist, bewirkt das Erdmagnetfeld eine Verringerung der lokalen Absorption, im Bereich der QT Näherung dagegen eine Vergrösserung. Nun ist im allgemeinen auf dem Lotungsweg im unteren Bereich die quasilongitudinale und in der Umgebung des Reflexionsniveaus die quasitransversale Näherung gültig. Das zeigt Fig. 3, die

* Bei dieser Definition bleiben die Kenngrössen für die Absorption der D-Schicht, B, und für die Absorption der E-Schicht, N, mit denen unseres früher benutzten Modells **B** vergleichbar. Abweichungen ergeben sich lediglich in den Fällen, in denen eine intensive E_s-Schicht auftritt.

VOL. 16 1959

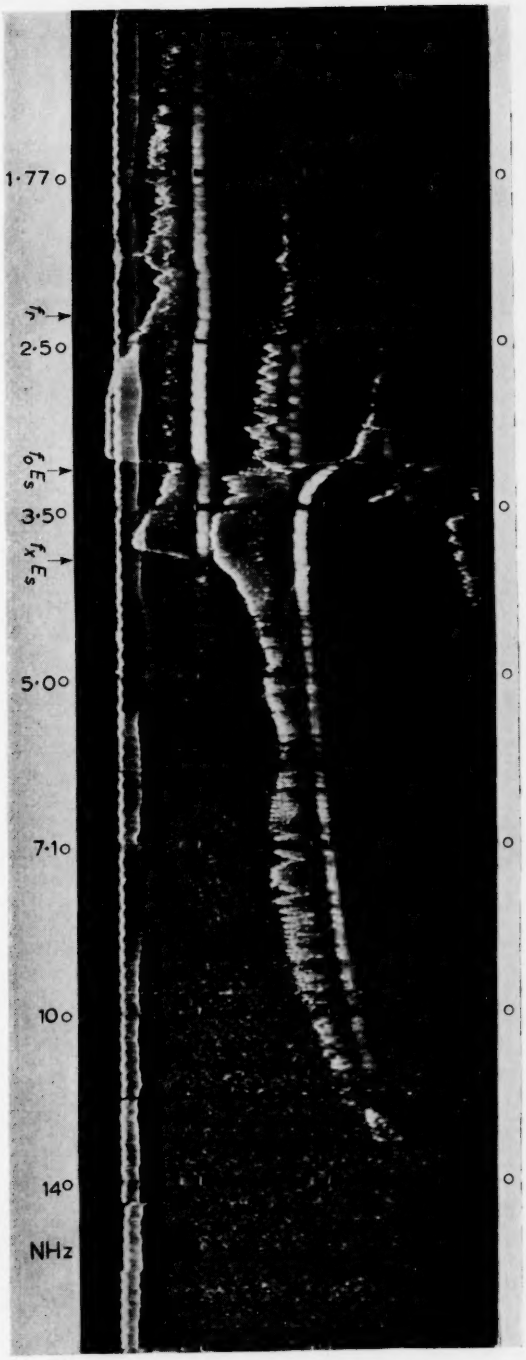


Fig. 1. Amplitudenregistrierung im Ionogramm.

VOL.
16
1959

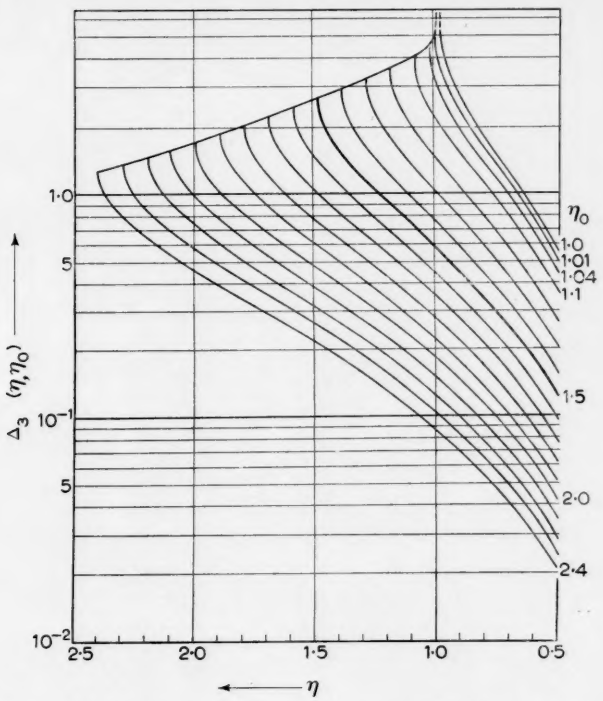


Fig. 2. Die Funktion $\Delta_3(\eta, \eta_0)$, Parameter η_0 (Beispiel $\eta_0 = 1{,}5$ fett).

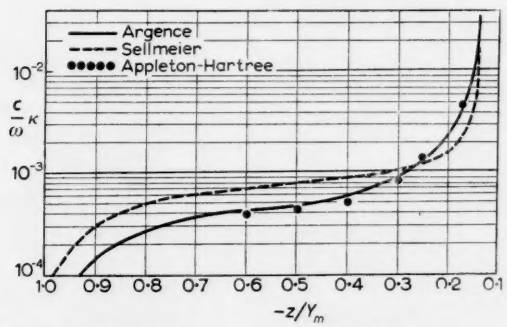


Fig. 3. Lokaler Absorptionskoeffizient für parabolische Schicht. ($f_0E = 3{,}5$ MHz, $f = 3{,}46$ MHz) (RAWER *et al.*, 1953).

durch numerische Berechnung erhalten wurde (RAWER *et al.*, 1953). Im Integral tritt eine gewisse Kompensation ein, und im Falle einer parabolischen Schicht ist diese Kompensation so gut, dass für den Integralwert der Absorption in mittleren Breiten nur ein geringer Einfluss des Erdmagnetfeldes besteht. Fig. 4 ist aus dem Zahlenmaterial von ARGENCE, RAWER und SUCHY entnommen und zeigt, dass die Abweichungen bei einer kritischen Frequenz von 3,5 MHz im Mittel bei 3% liegen.

Somit kann festgestellt werden, dass der Einfluss des Erdmagnetfeldes keine wesentliche Veränderung in unsere numerische Rechnung hereintragen kann, soweit sie die Reflexion an der E-Schicht betrifft. Etwas anders wird die Lage, wenn eine kräftig ionisierte E_s-Schicht auftritt, weil nun die Integration nicht mehr so tief in den QT Bereich erstreckt wird. Die Verschiebung der Integrationsgrenze kann dazu führen (vgl. Fig. 3), dass die Integration lediglich in dem Bereich erfolgt, in dem die Absorption durch den magnetischen Einfluss verringert wird. In diesen Fällen tritt keine Kompensation mehr ein, und wir haben zu erwarten, dass durch den Einfluss des Magnetfeldes die Absorption auf dem rechten Ast

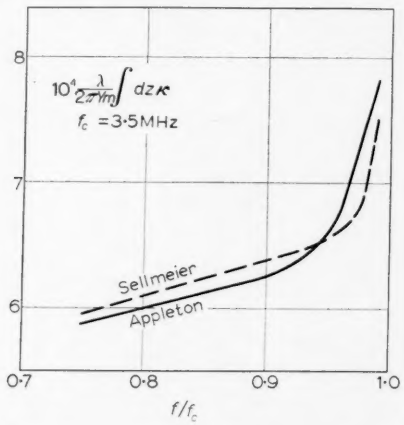


Fig. 4. Variation des Absorptionsintegrals mit der relativen Frequenz mit (Appleton) und ohne (Sellmeier) erdmagnetischen Einfluss.

mit fallenden noch stärker absinkt, als in Fig. 2 wiedergegeben. Schätzen wir nach Fig. 3 ab, so könnte dieser Einfluss in extremen Fällen eine Abnahme bis auf 3/4 des feldfreien Wertes bewirken.

4. Ergebnisse mit der neuen Methode

Unsere Absorptionsmessung wird seit 1957 auf günstig gestaffelten Frequenzen zwischen 1,725 und 5,75 MHz durchgeführt. Registriert wird der von einem logarithmischen Empfänger wiedergegebene Ausgangsimpuls, wobei je Frequenz über etwa 20 sec integriert wird, und jede der acht Frequenzen alle drei Minuten abgetastet wird.

Eine Beobachtungsperiode dauert zweimal 12 min. Mit einer geeichten Skala kann die Amplitude der Echos in jedem Stehbild abgelesen werden. In der für das AGI vorgesehenen Art und Weise wird durch Vergleich mit Nachtbeobachtungen das Absorptionsdekrement bestimmt.

Wir haben die Mittagsmessungen der Sommermonate 1958 (Messzeit 1145–1215 TU) unter Berücksichtigung der E_s-Reflexionen nach unserer neuen Theorie ausgewertet. Dabei muss nun natürlich auf jede einzelne Messperiode einzeln eingegangen werden. Die schon in (Bibl and Rawer, 1951) erwähnte "Spinnennetz-Methode" wurde nun mit Werten aus Fig. 2 angewandt. Es wurde also auf einer

VOL. 16 1959

B-Achse jeweils der Wert $\delta_a \cdot (f + f_H)^2$ und auf einer N-Achse der Wert δ_a/Δ_3 aufgetragen; beide Punkte wurden durch eine Gerade verbunden. Die Geraden für die verschiedenen Messgrössen sollten sich in einem Punkte schneiden. Sie tun das nicht exakt, aber doch in guter Näherung (Beispiel für alte und neue Methode in Fig. 5). Aus dem Zentralwert der verschiedenen Schnittpunkte bestimmt sich ein B-und ein N-Wert, also nicht-selektiver und selektiver Absorptionsanteil.

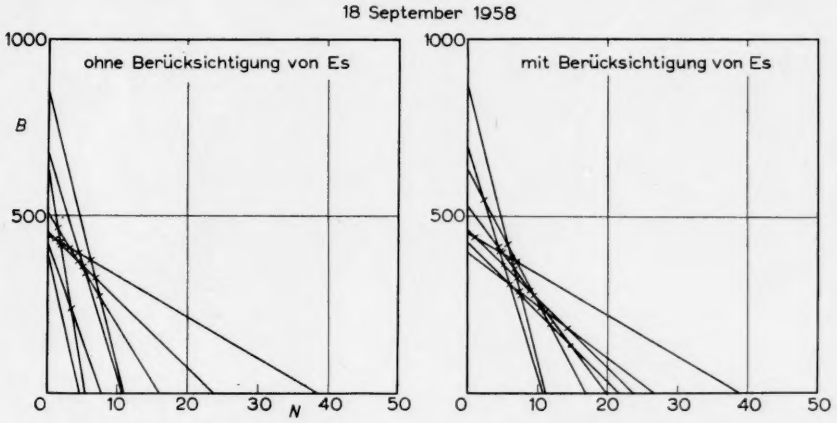


Fig. 5. Beispiel eines "Spinnennetz"-Diagramms, die gleiche Messung ausgewertet nach Modell **B** (links) und **D**(rechts).

Die Einzelergebnisse sind in Fig. 6 wiedergegeben, aus der zu ersehen ist, dass beide Charakteristiken keinen allzu grossen Schwankungen unterlagen. Die Gesamtstreubreite beträgt für B etwa $\pm 20\%$ und für N etwa $\pm 30\%$. Ob mehrtägige Perioden eine Rolle spielen, lässt sich aus den vorliegenden Ergebnissen noch nicht feststellen.

Besonders interessant ist der Fall eines Mögel-Dellinger Effekts (SID) am 29/4/1958. Fig. 7 zeigt das zugehörige "Spinnennetz", aus dem sich ein aussergewöhnlich hoher B-Wert von 620 ergibt, während der zugehörige N-Wert im normalen Streubereich liegt. Das stimmt gut mit der Vorstellung überein, dass der Effekt in erster Linie die D-Absorption beeinflusst.

Die jahreszeitliche Abhängigkeit der halbmonatlich gemittelten Ergebnisse zeigt Fig. 8 (Median-Werte und quartile). Für B wie für N besteht eine deutliche Abhängigkeit von der Jahreszeit, die sich jedoch als gegenläufig ergibt. Während B im Mittsommer ein leichtes Minimum zeigt, ergibt sich für N ein deutliches Maximum. Es wird notwendig sein, diese Jahresgänge in den Winter hinein zu verfolgen. Bei globaler Auswertung der Absorption hat sich bisher immer ein Maximum im Sommer ergeben; nach unserem Resultat sollte man annehmen, dass dieses Verhalten in erster Linie der selektiven Absorption zuzuschreiben ist. Bei unserer früheren Auftrennung nach Modell **B** hatte sich schon für N ein klarerer Jahresgang ergeben als für B. Die Fluktuation der damals erhaltenen Werte scheint grösser (Bibl und Rawer, 1951) als die jetzt erhaltene.

5. Vergleich verschiedener Ansätze

Um festzustellen, ob unsere Methode eine wirkliche Verbesserung bedeutet, haben wir eine grosse Zahl von Fällen nach den Ansätzen **A**, **B**, **C** und **D** dargestellt. Die Modell-Parameter wurden nach der üblichen Methode bestimmt, also für **A** als Median-Wert von $\delta_a \cdot (f + f_H)^2$, für **B**, **C** und **D** nach der Spinnennetz-Methode.

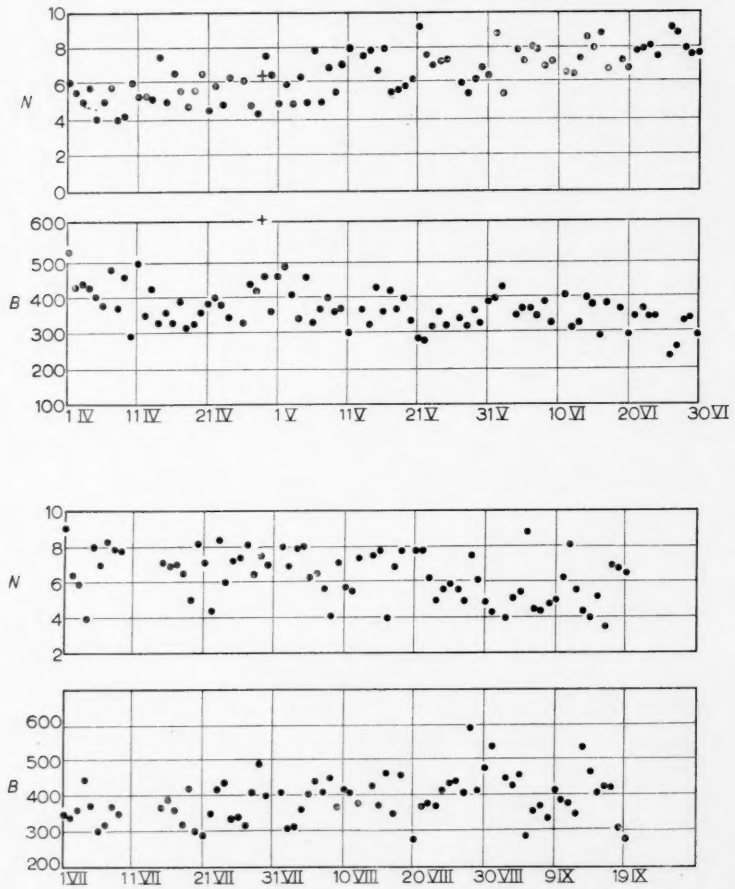


Fig. 6. Tägliche Mittagswerte für N und B. (+: SID)

Mit diesen Parametern wurde rückläufig für jede Frequenz ein theoretisches Dekrement L_{th} ausgerechnet, das mit dem beobachteten L_{obs} verglichen wurde. Die Differenz $\Delta L = L_{\text{th}} - L_{\text{obs}}$ wurde für den Monat April in Fig. 9 aufgetragen und zwar der quadratische Mittelwert aus allen acht Messfrequenzen. Man erkennt deutlich, dass Modell **A** die Messwerte erheblich schlechter wiedergibt als die drei anderen Modelle. Modell **C** ist gegen **A** eine sehr merkliche Verbesserung.* Modell **B** ist häufig ebensogut wie **D**, nämlich definitionsgemäss zu den

* Die nach Gl. (3) erhaltenen B_1-Werte wurden auch einzeln mit unseren B-Werten verglichen. Sie sind von gleicher Grössenordnung, jedoch ist die Strenung der B_1 erheblich stärker.

VOL. 16 1959

Zeiten, wo kein E_s auftrat. Sofern E_s auftrat, ist jedoch die Differenz von Modell **D** wesentlich geringer als die von Modell **B**.

In Fig. 10 sind Modell **B** und **D** für die feste Frequenz 4,08 MHz verglichen worden (Halbmonatsintervalle); auf dieser Frequenz tritt am häufigsten ein E_s-Echo auf. Die Überlegenheit von Modell **D** (obere Zeile) ist klar erkenntlich.

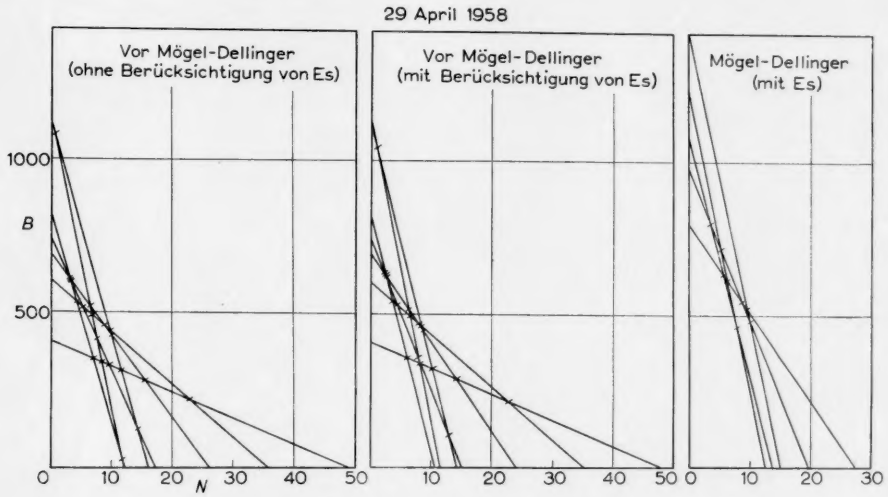


Fig. 7. "Spinnennetz"-Diagramm vor und während eines MD-Effekts.

Dasselbe ergibt sich beim Vergleich von Modell **A** and **D** in Fig. 11, in der jeweils die Frequenzabhängigkeit (Median und Quartils) aus 29 Apriltagen dargestellt ist. Die Fehler bei der neuen Methode liegen bei wenigen dB.

Schliesslich gibt Fig. 12 für alle acht Messfrequenzen in halbmonatlichen Abschnitten Median und Quartils des Darstellungsfehlers für Modell **D** wieder. Der mittlere Fehler überwiegt nur selten 5 dB, auch wenn man die verschiedenen Frequenzen einzeln betrachtet.

6. Ergebnisse

Unser neues Modell erfordert getrennte Betrachtung der Einzelmessungen, gibt aber dafür eine wesentlich befriedigendere Darstellung des Frequenzganges als bisher. Trotzdem bleibt es damit möglich, die Messergebnisse mit wenigen definierten Parametern wiederzugeben. Daraus ergibt sich die Aussicht, zu präziseren Vorhersagen der Absorption bei schrägem Einfall zu kommen.

Wenn früher vermutet wurde, dass die Schwankungen der Absorption in erster Linie der D-Region und damit B zuzuschreiben sind, so müssen wir jetzt feststellen, dass die Schwankungen der selektiven Absorption in der E-Region, nämlich die von N, meist noch grösser sind. Die physikalische Interpretation des letzten Resultats kann in zweifacher Weise geschehen, da in N das Produkt von Schichtdicke und Stosszahl (in Schichtmitte) eingeht. Die Auswertungen der Ionogramme zeigen, dass im Sommer die Höhe der Unterseite der E-Schicht niedriger liegt und die Schichtdicke etwas grösser ist. Beide Veränderungen

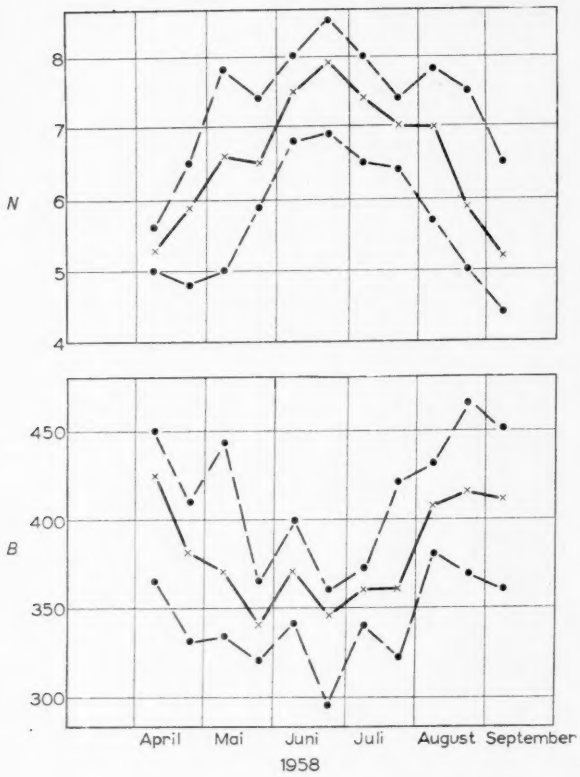


Fig. 8. Jahreszeitlicher Gang von B und N (Medianwert und Quartile).

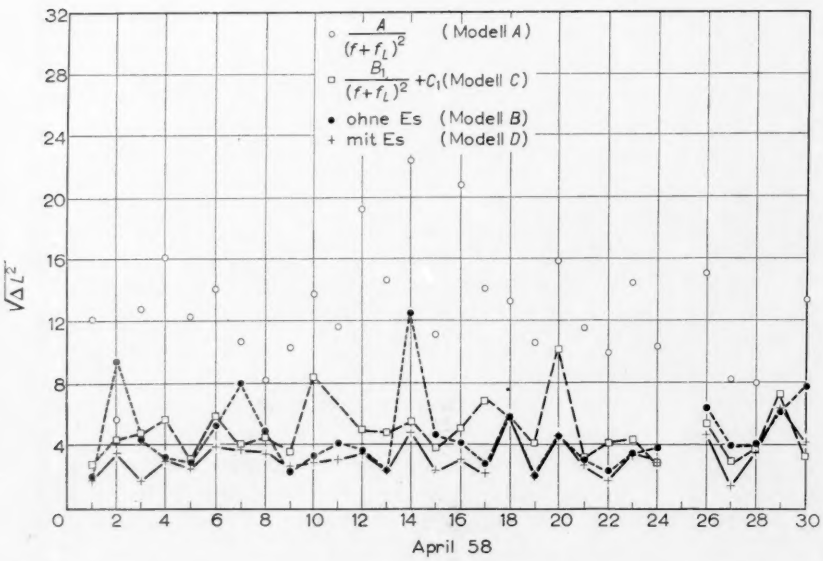


Fig. 9. Darstellungsfehler (quadratischer Mittelwert von 8 Frequenzen), t, **C** (—.—) und für die Modelle **A** (— —), **B** (. . .) **D** (— — —)

VOL. 16 1959

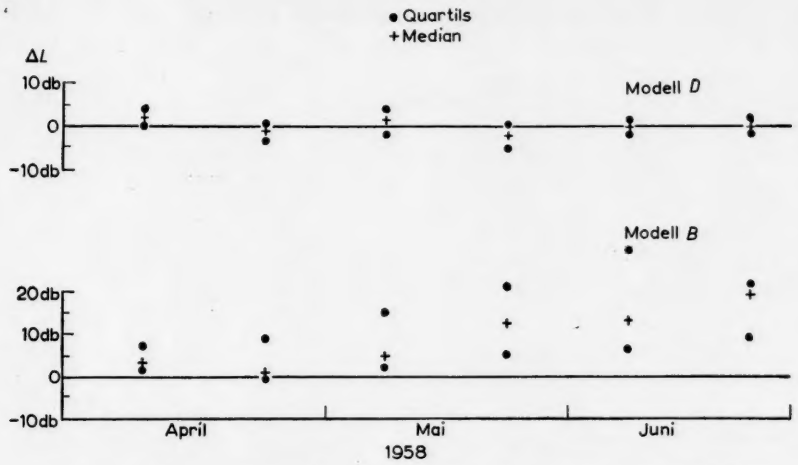


Fig. 10. Halbmonatliche Medianwerte des Darstellungsfehlers auf 4,08 MHz.

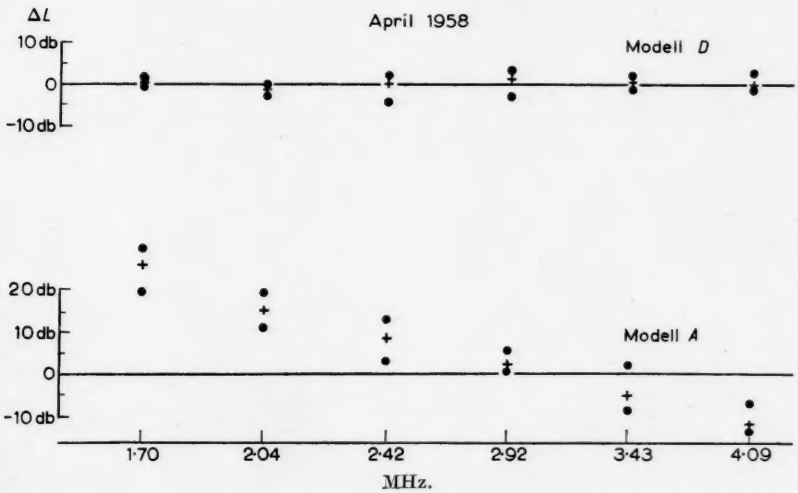


Fig. 11. Halbmonatliche Medianwerte des Darstellungsfehlers bei Modell **B** (unten) und **D** (oben), 8 Messfrequenzen.

können die beobachtete Zunahme von N erklären. Nehmen wir an, dass der Höheneinfluss überwiegt, so kann aus der Definitionsgleichung von N dei Stosszahl ν_0 bestimmt werden. In der folgenden Tabelle haben wir eine feste Halbdicke der E-Schicht ($Y_m = 17$ km) und eine Höhenvariation der Stosszahl nach $\exp(-2{,}5\ y)$ vorausgesetzt; das entspricht einer isothermen Skalenhöhe von 6,8 km in der E-Schicht. Aus den ν_0-Werten wurde die entsprechende Höhenschwankung ΔY berechnet. Die berechneten Schwankungen scheinen uns durchaus im Bereich des Möglichen zu liegen.

Die jahreszeitliche Veränderung von B ist umgekehrt als früher angenommen.

Wir finden im Mittsommer einen etwas niedrigeren Wert. Zur Klärung dieses Verhaltens sind weitere Untersuchungen und die Auswertung von Winterbeobachtungen erforderlich.

Interessant sind schliesslich noch die Absolutwerte von N und B. Der von N liegt im Mittel bei 6,6, im Sommer 1950 fanden wir 5,0 (Bibl and Rawer, 1951).

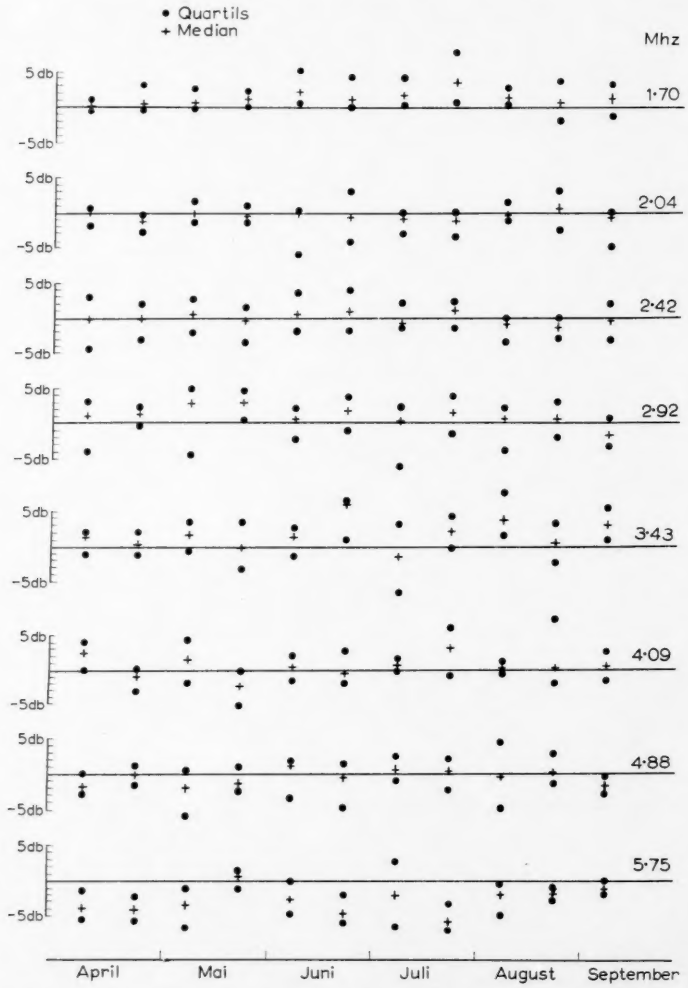


Fig. 12. Darstellungsfehler bei Modell **D**, Halbmonatswerte für die 8 Frequenzen.

Für B finden wir jetzt 370, damals 250. Hier ist zu bedenken, dass Methode D in Zweifelsfällen niedrigere N-Werte ergeben wird. Ausserdem haben wir den Einfluss der Sonnenfleckenrelativzahl R zu beachten, die im Sommer 1958 bei 190, damals aber bei 90 lag. Grob gesagt scheinen bei so hoher Aktivität beide Absorptionswerte etwa mit $\sqrt{R}$ zu variieren. In diesem Resultat sind natürlich

VOL. 16 1959

Tabelle 1.

Monat 1958	April		Mai		Juni		Juli		August		September
beobachtetes N (dB)	5,3	5,9	6,5	6,5	7,5	7,9	7,4	7,0	7,0	5,9	5,3
daraus: r_0 kHe)	10,9	12,1	13,5	13,3	15,4	16,2	15,2	14,4	14,4	12,1	10,9
Δy	0	−0,042	−0,086	−0,080	−0,138	−0,159	−0,133	−0,112	−0,112	−0,042	0
$Y_m \Delta y = \Delta Y$ (km)	0	−0,71	−1,46	−1,36	−2,35	−2,70	−2,26	−1,90	−1,90	−0,71	0

verschiedene Wirkungen des Zyklus enthalten, da sowohl die Ionisierung wie auch die Höhe der Schichten für den Absorptionswert von Bedeutung sind.

Anhang

Quasilongitudinale und quasitransversale Näherung für den Absorptionskoeffizienten κ

ω Pulsation — θ Winkel zwischen Magnetfeld und Ausbreitungsrichtung

ω_0 Plasma-Pulsation

μ_1 Realteil des Brechungsindex — $\sin^2\theta = \vartheta$

ν Stosszahl — ωH Gyro-Pulsation

c Lichtgeschwindigkeit — $X = \omega_0^2/\omega^2$

$Y = \omega_H/\omega$

Näherung nach der Appleton-Hartree-Formel für $\omega_0 \ll \omega$:

$$\frac{2c\mu_1}{\nu X}\kappa = \frac{1 + \frac{1}{2}Y^2\vartheta \mp \frac{1}{4}Y^4\vartheta^2/\sqrt{Y^2(1-\vartheta) + \frac{1}{4}Y^4\vartheta^2}}{1 - \frac{1}{2}Y^2\vartheta \pm \sqrt{Y^2(1-\vartheta) + \frac{1}{4}Y^4\vartheta^2}}. \tag{6}$$

1. *Quasilongitudinale Näherung* (QL)

Entwicklung nach Potenzen von ϑ ergibt:

$$\text{Zähler} = 1 + \tfrac{1}{2}Y^2\vartheta \mp \tfrac{1}{4}Y^3\vartheta^2 + O(\vartheta^3)$$
$$\text{Nenner} = (1 \pm Y)^2/\{1 \pm Y\vartheta + \tfrac{1}{4}Y(2Y \pm 1)\vartheta^2 + O(\vartheta^3)$$

Durch Ausmultiplizieren:

$$\frac{2c\mu_1}{\nu X}\kappa = \frac{1}{(1 \pm Y)^2}\left\{1 + \tfrac{1}{2}Y(Y \pm 2)\vartheta \pm \tfrac{1}{4}Y(1 \pm Y)^2\vartheta^2 + O(\vartheta^3)\right\} \tag{7}$$

Dies ist die QL-Näherung, gültig für kleines ϑ

Die Booker'sche Formel gibt statt dessen $\dfrac{1}{(1 \pm Y_L)^2}$

$$\omega_0 Y_L = Y \,.\, \cos\theta = Y \,.\, \sqrt{(1-\vartheta)}.$$

Daraus wurde durch entsprechende Entwicklung erhalten:

$$\frac{2c\mu_1}{\nu X}\kappa = \frac{1}{(1 \pm Y)^2}\left\{1 \pm \frac{Y}{1 \pm Y}\vartheta \pm \tfrac{1}{4}Y\frac{1 \pm 4Y}{(1 \pm Y)^2}\vartheta^2 + O(\vartheta^3)\right\} \tag{8}$$

Schon der Faktor erster Ordnung in ϑ in (8) unterscheidet sich merklich von dem in (7). Nur in nullter Ordnung stimmt (8) mit (7) überein, sodass der Ersatz von $(1 \pm Y)^{-2}$ durch $(1 \pm Y_L)^{-2}$ keine Verbesserung darstellt.

2. *Quasitransversale Näherung* (QT)

Sie könnte in analoger Weise bei kleinem $(1 - \vartheta)$ aus (6) erhalten werden. Jedoch wäre eine solche Formel praktisch uninteressant, weil die QT-Näherung nur in der Umgebung des Reflexionspunktes zutrifft (Rawer *et al.* 1953) und dort die Voraussetzung der Formel (6) $\omega_0 \ll \omega$ nicht zutrifft.

VOL. 16 1959

Literatur

Argence E., Mayot M. und Rawer K.	1950	*Ann. Géophys.* **6,** 242.
Beynon W. J. G. und Davies K.	1955	*The Physics of the Ionosphere* pp. 40–52. Physical Society, London.
Bibl K. und Rawer K.	1951	*J. Atmosph. Terr. Phys.* **2,** 51.
Booker H. G.	1935	*Proc. Roy. Soc.* A **150,** 267.
McAllcock G.	1955	*The Physics of the Ionosphere* pp. 14–19. Physical Society, London.
Piggott W. R.	1953	*Proc. Instn. Elect. Engrs.* Pt. 3, **100,** 61.
Piggott W. R., Beynon W. J. G. und Brown G. M.	1957	*Ann. IGY* **3,** Pt. 2, 179, 224.
Ratcliffe J. A.	1935	*Proc. Roy. Soc.* A **151,** 370.
Rawer K.	1951	*J. Atmosph. Terr. Phys.* **2,** 38.
Rawer K., Suchy K. und Argence E.	1953	*C.R. Acad. Sci., Paris* **236,** 190.
Whitehead J. D.	1957	*J. Atmosph. Terr. Phys.* **10,** 12.

L.
6
59

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 340 to 343. Pergamon Press Ltd. Printed in Northern Ireland

Some measurements of collision frequency in the *E*-region of the ionosphere

D. M. SCHLAPP
Cavendish Laboratory, Cambridge*

(*Received* 12 *February* 1959)

Abstract—Measurements of collision frequency in the *E*-region were made by observing how the deviative absorption of a radio echo varied with its group path as a critical frequency was approached. The collision frequency at heights between 105 km and 120 km was found to vary with a scale height of about 16 km. In 1955, the collision frequency passed through a value of 2×10^4 sec^{-1} at a height of about 112 km. There is some evidence that, when the sunspot number is greater, the collision frequency at a fixed height is greater.

1. INTRODUCTION

MANY measurements (FARMER, 1936; BRIGGS, 1951a) of the electron collision frequency (ν) in the *E*-region have been made in the past by the method of comparing the change in reflection coefficient ρ, and the corresponding changes in group path (P') and phase path (P), resulting from a small change in the frequency of a pulsed ionospheric sounding apparatus. The expression

$$\Delta \log \rho = -\frac{\nu}{2c}(\Delta P' - \Delta P)$$

due to APPLETON (1935) was used to calculate the average value of ν over the range of height responsible for the changes. In none of these measurements, however, was the height of the *E*-layer known accurately, and it is therefore not easy to compare the results with what would be expected theoretically.

The purpose of the investigation described in the present paper was to make measurements of the collision frequency at different heights in the *E*-region with the greatest possible accuracy, and under conditions such that the height to which the measurements referred could be deduced. The measurements were made with the help of an apparatus, described by BRIGGS (1951b), in which the necessary frequency change could be made in 1 or 2 sec, and the corresponding changes of amplitude and echo delay were exhibited on Cartesian co-ordinate axes on a cathode-ray tube. Many readings of amplitude were taken for each determination of collision frequency in order to average out the effects of fading. Obviously disturbed conditions were avoided.

The h' (f) curve was recorded at the time of making the measurements, and the real height of the peak of electron density in the *E*-region was calculated from it by a method which included the effect of the earth's magnetic field. It is considered that the height was estimated with an accuracy of ± 3 km. There

* Now at University College of North Staffordshire.

VOL. 16 1959

may also be a systematic error due to low lying electrons with plasma frequency below the lower limit (1·5 Mc/s) of the ionosonde used; this error is probably less than 5 km.

2. Results

In order to investigate how the collision frequency varied with height, all measurements of collision frequency were grouped according to height in 5 km ranges. The mean values are shown plotted in Fig. 1, in which the scale of collision frequency is logarithmic. The measurements have been divided into those obtained in 1955 and in 1956. It is seen that there was a significant increase in the collision frequency at a given height from 1955 to 1956.

The collision frequency deduced for a fixed height shows no significant change either with time of day or year.

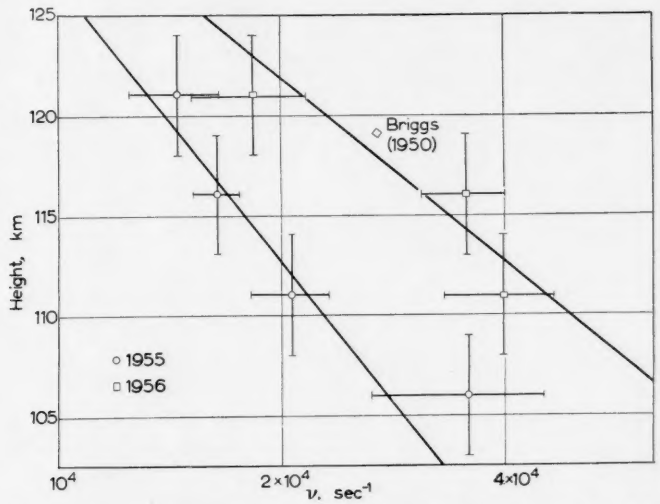


Fig. 1. Variation of collision frequency with height in the E-region. The lines indicate the standard deviations of the measurements.

An attempt was made to compare values obtained from measurements at frequencies just below a critical frequency with values obtained on the same occasion from measurements on the F- or E_s- layer echo at frequencies just above the same critical frequency. The difference between the means of such sets of values was not significant. Only a small number of results could be obtained, however, since conditions were rarely suitable for measurement on both sides of a critical frequency simultaneously.

An attempt was also made to compare pairs of values of collision frequency obtained from measurements made on the penetration of the E-layer by the ordinary and extraordinary waves. For this purpose the expression

$$\Delta \log \rho = -\frac{\nu}{2c} \frac{1}{1 - Y_L/2} (\Delta P' - \Delta P)$$

was used for the extraordinary wave. Since the extraordinary wave is highly absorbed at frequencies which are reflected from the normal E-layer, the measurement was possible only when strong sporadic E-reflections were present on a higher frequency. The difference between the means of the values obtained from measurements on the two component waves was not significant, but the number of results was small.

3. DISCUSSION

The values obtained in the present series of experiments agree well in general with values obtained previously by other workers, and in particular, they agree with some results obtained from absolute measurements of absorption. For instance, BEYNON and DAVIES (1955) deduced a value of $3{\cdot}4 \times 10^4$ sec^{-1} at 110 km from measurements of absorption at 2 Mc/s.

If the results of Fig. 1 are expressed in the form $\nu = \nu_0 e^{-h/H}$ the scale height H is about 20 km for 1955 and about 13 km for 1956. Although no great accuracy is claimed for these values, it seems that at least for 1955, the scale height was a good deal greater than that expected by NICOLET (1953) whose curve corresponds to $H \fallingdotseq 10$ km at the levels considered here. The observed collision frequencies agree with NICOLET's theoretical curve at about 110 km, but are greater at the greater heights.

According to KALLMANN *et al.* (1956), oxygen starts to dissociate at a height of about 90 km and is 75–85% dissociated at 130 km. Some workers (BATES and MASSEY, 1948; YAMANOUCHI, 1947) have suggested that the cross-section for collisions between electrons and oxygen atoms is very large, corresponding to a resonance effect. This might explain why the observed values of collision frequency at the greater heights are larger than those predicted by NICOLET who assumes a small cross-section for collisions between electrons and oxygen atoms.

The mean Zürich sunspot number for the period from May to December 1955 was 49, and for January to July 1956 it was 113. It is possible that the difference in the collision frequencies measured during these two periods is related to this increase.

It may be relevant that in 1950 when the mean Zürich sunspot number was 84, BRIGGS (1951a) found a mean value of $2{\cdot}7 \times 10^4$ sec^{-1} at a mean height of 119 km. This point is shown in Fig. 1 for comparison.

4. CONCLUSIONS

The collision frequency at heights between 105 km and 120 km varies with a scale height of about 16 km and in 1955 passed through a value of 2×10^4 sec^{-1} at a height of about 112 km.

There is some evidence that when the sunspot number is greater, the collision frequency at a fixed height is greater.

Acknowledgements—The writer would like to express his gratitude to Mr. J. A. RATCLIFFE and Dr. B. H. BRIGGS for the advice given in many helpful discussions. He is indebted to the Department of Scientific and Industrial Research for the award of a maintenance grant.

VOL. 16 1959

References

Appleton E. V.	1935	*Nature, Lond.* **135,** 618.
Bates D. R. and Massey H. S. W.	1948	*Proc. Roy. Soc.* A **192,** 1.
Beynon W. J. G. and Davies K.	1955	*The Physics of the Ionosphere* p. 40. Physical Society, London.
Briggs B. H.	1951a	Thesis, University of Cambridge.
Briggs B. H.	1951b	*Proc. Phys. Soc. Lond.* B **64,** 255.
Farmer F. T.	1936	Thesis, University of Cambridge.
Kallmann H. K., White W. B. and Newell H. E.	1956	*J. Geophys. Res.* **61,** 513.
Nicolet M.	1953	*J. Atmosph. Terr. Phys.* **3,** 200.
Yamanouchi T.	1947	*Progr. Theor. Phys., Osaka* **2,** 33.

OL.
16
959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 344 to 356. Pergamon Press Ltd. Printed in Northern Ireland

A theory for determining upper-atmosphere winds from radio observations on meteor trails

G. V. GROVES
Department of Physics, University College, London

(*Received* 26 *February* 1959)

Abstract—Least-squares theory is applied to observations on drifting meteor trails to obtain the determination of a general number of parameters, characterizing the wind structure. The results enable vertical air motion and time- and height-variations in the wind structure to be examined, whereas previously a constant horizontal wind was assumed. It is shown how, on the basis of the least-squares solution, the errors in the parameters, and hence in the calculated wind structure, can be estimated. The method is applied to a typical set of data.

1. INTRODUCTION

THE first determinations of wind velocity in the region of 90 km altitude by means of radio reflections from meteor trails were undertaken at Stanford University by MANNING *et al.* (1950). Since that time, the technique has been brought into regular use at Adelaide and Mawson by the Department of Physics, University of Adelaide, and at the Jodrell Bank Experimental Station, University of Manchester.

VOL. 16 1959

In the Adelaide method, C.W. radiation is emitted vertically within a cone of semi-angle 40°, and a meteor trail suitably orientated within this cone reflects radiation back to an aerial array on the ground, by means of which the direction of the reflection point on the trail can be determined. The component velocity of drift in the direction of observation is derived from the doppler frequency change of the reflected wave, while the range of the reflection point is found by the radar-technique using pulses superimposed on the C.W. A fuller description of the method appears in the *IGY Instruction Manual* (BEYNON and BROWN, 1957). For a single trail the following quantities are measured:

l, m, n, the direction cosines of the reflection point on the trail
V, the component of drift velocity in the direction of observation
z, the height of the reflection point
t, the time of observation.

Although the Manchester method differs considerably in experimental technique, it leads to the same set of observational data as the Adelaide method. No attempt is made to measure the direction of the reflection point, and so l, m and n have to be taken as the direction of the axis of the illuminating beam and are subject to errors corresponding to a beam width of $\pm 30°$. On the other hand, the echo rate is some ten to a hundred times greater and this compensates for the uncertainty in direction. In order to obtain the total wind vector the aerial is pointed alternately in two directions at right angles for short periods (BEYNON and BROWN, 1957).

The method of analysis which has been applied to the Adelaide data is that used by MANNING *et al.* (1950). In this, observations are taken, which occur within sufficiently short time- and height-ranges for the wind variations to be neglected.

A wind amplitude and a wind direction (assumed horizontal) are then determined, which fit the observations according to the method of least squares. Many groups are analysed in this way for different time- and height-ranges; and the NS and EW components obtained from the least-squares fits are then analysed for 24 hr and 12 hr periodicities and sometimes for altitude variations. The Manchester analysis follows the same procedure; although, with this, only a simple averaging process is needed to obtain the wind vector since all echoes within a given aerial beam have the same direction assigned to them.

In theory, the above procedure would be correct if a large number of trails could be observed within infinitesimally small intervals of time and height. In practice, the situation often departs considerably from the ideal. On account of the limited number of observations and the need to smooth random variations, the time and height intervals have to be increased, in conflict with the assumption that the wind variations within these intervals are negligibly small. The situation becomes rapidly more unfavourable for height intervals further away from the level of maximum meteor density.

The present paper removes the assumption that variations in wind are negligible within certain time and height intervals by generalizing the least-squares solution so that parameters defining the time- and height-variations of the wind may be determined. The solution should enable more information of greater significance to be obtained from a given quantity of data. The determination of the wind variation with height, in particular, is expected to be improved.

This paper was written while the writer was a visitor to the Department of Physics, University of Adelaide, and the discussions which he had with Dr. W. G. Elford contributed substantially to his understanding of the meteor wind-measuring technique.

2. General Formulation of Least-squares Solution

Let u, v and w be the three components of wind velocity at the reflection point; then the velocity V of the atmosphere in the direction (l, m, n) is

$$V = ul + vm + wn \tag{1}$$

V is the measured doppler velocity, provided the trail moves parallel to itself while the velocity determination is being made. Although wind gradients produce a rotation of the trail, the effect on the doppler velocity would be expected to be very small since it depends on the angle of rotation to the second order. Also, no detectable effect has been apparent in the doppler records. Hence equation (1) will be taken to be valid, when V is the measured doppler velocity.

The experimental determination of u, v and w is undertaken on the assumption that they can be expressed as certain specified functions of z and t

$$\begin{aligned} u &= u(z, t, a_{1n_1}) \\ v &= v(z, t, a_{2n_2}) \\ w &= w(z, t, a_{3n_3}) \end{aligned} \tag{2}$$

where a_{kn_k} $(n_k = 1, \ldots, N_k;\ k = 1, 2, 3)$ are parameters whose determination is required.

On account of random fluctuations in (a) the experimental measurements, and (b) the wind structure, the quantities

$$\varepsilon_i = u_i l_i + v_i m_i + w_i n_i - V_i \tag{3}$$

where the suffix i refers to the ith trail and $u_i = u(z_i, t_i, a_{1n_1})$, $v_i = v(z_i, t_i, a_{2n_2})$, $w_i = w(z_i, t_i, a_{3n_3})$, ($a_{kn_k}$ having suitably determined values), would be expected to be random fluctuations with zero mean. The distribution of such random fluctuations is usually nearly normal in physical phenomena because they arise from a multitude of minor random fluctuations. Accordingly, it is assumed that ε_i are independently and normally distributed with zero mean and s.d. $\sigma\varepsilon_i$ for $i = 1, \ldots M$, the number of meteors observed being M.

The probability that the values of ε_i for $i = 1, \ldots M$ should be in the ranges $(\varepsilon_1, \varepsilon_1 + d\varepsilon_1) \ldots (\varepsilon_m, \varepsilon_m + d\varepsilon_m)$, respectively is then

$$P(\varepsilon_1, \ldots, \varepsilon_m)\, d\varepsilon_1 \ldots d\varepsilon_m = \prod_{i=1}^{M} \frac{e^{-\varepsilon_i^2/2\sigma^2\varepsilon_i}}{\sigma\varepsilon_i\sqrt{(2\pi)}}\, d\varepsilon_i \tag{4}$$

The parameters a_{kn_k} will be chosen so that the observed set of values $\varepsilon_1, \ldots \varepsilon_M$ has maximum probability, i.e. so that $\sum_{i=1}^{M} \varepsilon_i^2/\sigma^2\varepsilon_i$ is a minimum with respect to a_{kn_k}. This is the least-squares method of solution and gives the conditions

$$\sum_{i=1}^{M} \omega_i^2 \varepsilon_i \partial\varepsilon_i/\partial a_{kn_k} = 0 \qquad (n_k = 1 \ldots N_k;\ k = 1, 2, 3) \tag{5}$$

where the weighting factors $\omega_i \propto 1/\sigma\varepsilon_i$. Equation (5) gives $N_1 + N_2 + N_3$ equations for a_{kn_k}.

VOL. 16 1959

3. Solution for Parameters a_{kn_k}

It is now assumed that equations (2) are linear in a_{kn_k}. Then (3) can be written as

$$\begin{aligned} \varepsilon_i &= \varepsilon_{i0} + \mathbf{a}[\partial\varepsilon_i/\partial a_{11} \ldots \partial\varepsilon_i/\partial a_{3N_3}]' \\ &= \varepsilon_{i0} + [\partial\varepsilon_i/\partial a_{11} \ldots \partial\varepsilon_i/\partial a_{3N_3}]\mathbf{a}', \end{aligned} \tag{6}$$

where

$$\begin{aligned} \varepsilon_{i0} &= u(z_i, t_i, 0)\, l_i + v(z_i, t_i, 0) m_i + w(z_i, t_i, 0) n_i - V_i \\ &= -V_i \quad \text{(See footnote*)} \end{aligned} \tag{7}$$

$$\mathbf{a} = [a_{11} \ldots a_{3N_3}] \tag{8}$$

By (6), (5) becomes

$$\mathbf{S}\boldsymbol{\epsilon}'_{0\omega} + \mathbf{S}\mathbf{S}'\mathbf{a}' = 0 \tag{9}$$

where

$$\mathbf{S} = \begin{bmatrix} \omega_1 \partial\varepsilon_1/\partial a_{11} \ldots \omega_M \partial\varepsilon_M/\partial a_{11} \\ \vdots \\ \omega_1 \partial\varepsilon_1/\partial a_{3N_3} \ldots \omega_M \partial\varepsilon_M/\partial a_{3N_3} \end{bmatrix} \tag{10}$$

$$= \begin{bmatrix} \omega_1 l_1 \partial u_1/\partial a_{11} \ldots \omega_M l_M \partial u_M/\partial a_{11} \\ \vdots \\ \omega_1 n_1 \partial w_1/\partial a_{3N_3} \ldots \omega_M n_M \partial w_M/\partial a_{3N_3} \end{bmatrix} \tag{11}$$

* $u(z_i, t_i, 0)$, $v(z_i, t_i, 0)$, $w(z_i, t_i, 0)$ can be taken as zero without loss of generality, unless numerical values of wind velocity are being incorporated into (2).

by (3), and

$$\boldsymbol{\epsilon}_{0\omega} = [\omega_1\varepsilon_{10} \ldots \omega_M\varepsilon_{M0}] = -[\omega_1 V_1 \ldots \omega_M V_M]. \tag{12}$$

The solution of (9) is*

$$\mathbf{a} = [\omega_1 V_1 \ldots \omega_M V_M]\,\mathbf{T} \tag{13}$$

where

$$\mathbf{T} = \mathbf{S}'(\mathbf{S}\mathbf{S}')^{-1} \tag{14}$$

4. The Weighting Factors, ω_i

If the distribution of ε_i were known, ω_i could be taken to be any set of values proportional to $1/\sigma\varepsilon_i$. In the first instance, $\sigma\varepsilon_i$ are unknown and some criterion needs to be introduced by which ω_i may be determined. The following is suggested. Assume that: (a) l_i, m_i, n_i, V_i, z_i, t_i are perfectly accurate, (b) a_{kn_k} are independently distributed, (c) a_{kn_k} have the same s.d.; (a) is a reasonable assumption, if we expect the residuals ε_i to arise from wind fluctuations and not from measurement errors, (b) holds approximately if the number of observations is large and they are randomly distributed, (c) can always be satisfied by introducing appropriate numerical factors into the coefficients of the a_{kn_k}'s (in equation (6)). Then from (3)

$$\sigma^2\varepsilon_i = \Big[l_i^2 \sum_{n_1} (\partial u_i/\partial a_{1n_1})^2 + m_i^2 \sum_{n_2} (\partial v_i/\partial a_{2n_2})^2 + n_i^2 \sum_{n_3} (\partial w_i/\partial a_{3n_3})^2\Big]\sigma^2 a$$

and we can take

$$\omega_i = \Big[l_i^2 \sum_{n_1} (\partial u_i/\partial a_{1n_1})^2 + m_i^2 \sum_{n_2} (\partial v_i/\partial a_{2n_2})^2 + n_i^2 \sum_{n_3} (\partial w_i/\partial a_{3n_3})^2\Big]^{-1/2} \tag{16}$$

This choice of ω_i can be given a simple geometrical interpretation. In the $N_1 + N_2 + N_2$–dimensional space with co-ordinates a_{kn_k}, the length of the perpendicular from a point $\hat{a}_{kn_k}$ on to the plane $\varepsilon_i\,(a_{kn_k}) = 0$ is

$$p_i = \omega_i\varepsilon_i(\hat{a}_{kn_k})$$

where ω_i is given by (16). Hence the least-squares condition (5) can be interpreted as minimizing the sum of the squares of the perpendicular lengths $\sum_i p_i^2$

5. All Observations Occur within a Short Time, All Reflection Points at Nearly the Same Height, and Vertical Air Motion Small

In the following sections, the theory will be applied to particular forms of (2). In the present case, (2) may be taken, independent of t and z, as

$$u = u_0$$
$$v = v_0$$

From (16)

$$\omega_i = 1/(l_i^2 + m_i^2)^{1/2} \tag{17}$$

From (11)

$$\mathbf{S} = \begin{bmatrix} l_1/(l_1^2 + m_1^2)^{1/2} \ldots l_M/(l_M^2 + m_M^2)^{1/2} \\ m_1/(l_1^2 + m_1^2)^{1/2} \ldots m_M/(l_M^2 + m_M^2)^{1/2} \end{bmatrix} \tag{18}$$

* Also

$$[\omega_1\varepsilon_1 \ldots \omega_M\varepsilon_M] = [\omega_1 V_1 \ldots \omega_M V_M]\,(\mathbf{TS} - \mathbf{I}), \tag{15}$$

where $\mathbf{I}$ is the unit matrix of order M. When $M = N_1 + N_2 + N_3$, $\mathbf{S}$ is square and $\mathbf{T} = \mathbf{S}^{-1}$, $\omega_i\varepsilon_i = 0$.

Hence

$$\mathbf{SS}' = \begin{bmatrix} \sum l^2/(l^2+m^2) & \sum lm/(l^2+m^2) \\ \sum lm/(l^2+m^2) & \sum m^2/(l^2+m^2) \end{bmatrix}$$

on omitting suffix i; then

$$(\mathbf{SS}')^{-1} = \begin{bmatrix} \sum m^2/(l^2+m^2) & -\sum lm/(l^2+m^2) \\ -\sum lm/(l^2+m^2) & \sum l^2/(l^2+m^2) \end{bmatrix} \div \left[\sum \frac{l^2}{l^2+m^2} \sum \frac{m^2}{l^2+m^2} - \left(\sum \frac{lm}{l^2+m^2} \right)^2 \right] \tag{19}$$

From (13), (14), (17) and (18)

$$[u_0\ v_0] = \left[\sum \frac{lV}{l^2+m^2} \sum \frac{mV}{l^2+m^2} \right] (\mathbf{SS}')^{-1}$$

Hence

$$u_0 = \left[\sum \frac{lV}{l^2+m^2} \sum \frac{m^2}{l^2+m^2} - \sum \frac{mV}{l^2+m^2} \sum \frac{lm}{l^2+m^2} \right] \div \left[\sum \frac{l^2}{l^2+m^2} \sum \frac{m^2}{l^2+m^2} - \left(\sum \frac{lm}{l^2+m^2} \right)^2 \right] \tag{20}$$

$$v_0 = \left[\sum \frac{mV}{l^2+m^2} \sum \frac{l^2}{l^2+m^2} - \sum \frac{lV}{l^2+m^2} \sum \frac{lm}{l^2+m^2} \right] \div \left[\sum \frac{l^2}{l^2+m^2} \sum \frac{m^2}{l^2+m^2} - \left(\sum \frac{lm}{l^2+m^2} \right)^2 \right] \tag{21}$$

If we write $l = \sin\gamma\cos\theta$, $m = \sin\gamma\sin\theta$, $V = V_\theta \sin\gamma$, V_θ being the horizontal wind component in the θ-azimuth, then (20) and (21) become

$$u_0 = (\sum s^2\ \sum cV_\theta - \sum sc\ \sum sV_\theta)/[\sum s^2 \sum c^2 - (\sum sc)^2]$$

$$v_0 = (\sum c^2\ \sum sV_\theta - \sum sc\ \sum cV_\theta)/[\sum s^2 \sum c^2 - (\sum sc)^2]$$

where $s = \sin\theta$, $c = \cos\theta$. Hence the wind direction and amplitude are

$$\tan\phi = v_0/u_0 = [\Sigma c^2 \Sigma sV_\theta - \Sigma sc \Sigma cV_\theta)/(\Sigma s^2 \Sigma cV_\theta - \Sigma sc \Sigma sV_\theta) \tag{22}$$

$$W = |u_0|\ (1 + \tan^2\phi)^{1/2} = (1 + \tan^2\phi)^{1/2}\ \Sigma cV_\theta/(\Sigma c^2 + \tan\phi\ \Sigma sc) \tag{23}$$

Equations (22) and (23) agree with the original analysis of MANNING *et al.* (1950).

The standard deviations σu_0, σv_0 are obtained from equation (A.6) in the appendix. In the case of a large number of randomly distributed trails, $\Sigma lm/(l^2+m^2) \sim 0$ and $\Sigma l^2/(l^2+m^2)$, $\Sigma m^2/(l^2+m^2) \sim M/2$. Hence

$$(\mathbf{SS}')^{-1} \sim \frac{2}{M} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and (A.6) gives

$$\sigma u_0,\ \sigma v_0 \sim (2/M)^{1/2}\ \sigma\varepsilon_\omega$$

Taking $\sigma\varepsilon_\omega = 25$ m/sec and $M = 60$, σu_0, $\sigma v_0 = 4{\cdot}6$ m/sec.

VOL. 16 1959

6. All Observations Occur within a Short Time

The equations (2) are taken to be independent of t and expressed in the form

$$\begin{aligned} u &= u_0 + u_1 S \ldots + u_A S^A \\ v &= v_0 + v_1 S \ldots + v_B S^B \\ w &= w_0 + w_1 S \ldots + w_C S^C \end{aligned} \tag{24}$$

where S is defined as

$$S = \frac{2z - (z_\beta + z_\alpha)}{z_\beta - z_\alpha},$$

z_β, z_α being the greatest and least of the values $z_1 \ldots z_M$. (Then $S = 0$ is the mid-point of the interval (z_β, z_α); and S has the values ± 1 at the end points. S is dimensionless and will be more convenient than z in numerical work.) From (16)

$$\begin{aligned} \omega_i = [l_i^2(1 + S_i^2 \ldots + S_i^{2A}) + m_i^2(1 + S_i^2 \ldots + S_i^{2B}) \\ + n_i^2(1 + S_i^2 \ldots + S_i^{2C})]^{-1/2} \end{aligned} \tag{25}$$

ω_i lies between 1 and $1/[\text{max of (A + 1), (B + 1), (C + 1)}]^{1/2}$. If the wind structure were taken to be linear $A = B = C = 1$, ω_i lies in $(1/\sqrt{2}, 1)$.

From (11)

$$\mathbf{S} = \begin{bmatrix} \omega_1 l_1 & \ldots \omega_M l_M \\ \vdots & \vdots \\ \omega_1 l_1 S_1^A & \ldots \omega_M l_M S_M^A \\ \omega_1 m_1 & \ldots \omega_M m_M \\ \vdots & \vdots \\ \omega_1 m_1 S_1^B & \ldots \omega_M m_M S_M^B \\ \omega_1 n_1 & \ldots \omega_M n_M \\ \vdots & \vdots \\ \omega_1 n_1 S_1^C & \ldots \omega_M n_M S_M^C \end{bmatrix} \tag{26}$$

All elements of $\mathbf{S}$ are $\leqslant 1$ in absolute value. Equation (9) needs to be solved numerically for

$$\mathbf{a} = [u_0 \ldots u_A \; v_0 \ldots v_B \; w_0 \ldots w_C]$$

In Section 8, a typical set of data is analysed, taking u, v to be quadratic and w as constant.

7. Height- and Time-variations Taken into Account

The equations (2) are taken to be periodic in t with a period of $2\pi/\omega$ and (on taking a finite number of terms of the Fourier expansion) expressed as

$$\begin{aligned} u &= u_0(z) + u_1(z) \sin \omega t + \ldots + u_p(z) \sin p\omega t \\ &\quad + u_1^*(z) \cos \omega t + \ldots + u_p^*(z) \cos p\omega t \\ v &= v_0(z) + v_1(z) \sin \omega t + \ldots + v_q(z) \sin q\omega t \\ &\quad + v_1^*(z) \cos \omega t + \ldots + v_q^*(z) \cos q\omega t \\ w &= w_0(z) + w_1(z) \sin \omega t + \ldots + w_r(z) \sin r\omega t \\ &\quad + w_1^*(z) \cos \omega t + \ldots + w_r^*(z) \cos r\omega t \end{aligned} \tag{27}$$

where

$$\begin{aligned}
u_j(z) &= u_{j0} + u_{j1}S \ldots + u_{jA_j}S^{A} && (j = 0 \ldots p)\\
u_j^*(z) &= u_{j0}^* + u_{j1}^*S \ldots + u_{jA_j}^*S^{A_j} && (j = 1 \ldots p)\\
v_j(z) &= v_{j0} + v_{j1}S \ldots + v_{jB_j}S^{B_j} && (j = 0 \ldots q)\\
v_j^*(z) &= r_{j0}^* + v_{j1}^*S \ldots + v_{jB_j}^*S^{B} && (j = 1 \ldots q)\\
w_j(z) &= w_{j0} + w_{j1}S \ldots + w_{jC_j}S^{C_j} && (j = 0 \ldots r)\\
w_j^*(z) &= w_{j0}^* + w_{j1}^*S \ldots + w_{jC_j}^*S^{C_j} && (j = 1 \ldots r)
\end{aligned} \tag{28}$$

From (16)

$$\begin{aligned}
\omega_i = [l_i^2 \sum_{j=0}^{p} (1 + S_i^2 \ldots + S_i^{2A_j}) + m_i^2 \sum_{j=0}^{r} (1 + S_i^2 \ldots + S_i^{2B_j})\\
+ n_i^2 \sum_{j=0}^{r} (1 + S_i^2 \ldots + S_i^{2C_j})]^{-1/2}
\end{aligned} \tag{29}$$

It is seen that

$$\begin{aligned}
[\max \text{ of } \sum_{j=0}^{p} (A_j + 1), \sum_{j=0}^{q} (B_j + 1), \sum_{j=0}^{q} (C_j + 1)]^{-1/2} \leqslant \omega_i\\
\leqslant [\min (p + 1), (q + 1), (r + 1)]^{-1/2} \leqslant 1
\end{aligned}$$

If $A_j = B_j = C_j = C$, $p = q = r$, then ω_i can be taken as $(1 + \ldots S_i^{2C})^{-1/2}$, being arbitrary as regards a constant multiplier. For $A_j = B_j = C_j = 1$, $p = q = r$, ω_i lies in $(1/\sqrt{2}, 1)$. From (11)

$$\mathbf{S} = \begin{bmatrix} \{\mathbf{L}_{jk}\} \\ \{\mathbf{M}_{jk}\} \\ \{\mathbf{N}_{jk}\} \end{bmatrix}, \tag{30}$$

where

$$\begin{aligned}
\{\mathbf{L}_{jk}\} &= [\mathbf{L}_{00} \ldots \mathbf{L}_{0A_0} \ldots \mathbf{L}_{p0} \ldots \mathbf{L}_{pA_p}]'\\
\{\mathbf{M}_{jk}\} &= [\mathbf{M}_{00} \ldots \mathbf{M}_{0B_0} \ldots \mathbf{M}_{q0} \ldots \mathbf{M}_{qB_q}]'\\
\{\mathbf{N}_{jk}\} &= [\mathbf{N}_{00} \ldots \mathbf{N}_{0C_0} \ldots \mathbf{N}_{r0} \ldots \mathbf{N}_{rC_r}]'\\
\mathbf{L}_{0k} &= [\omega_1 l_1 S_1^k \ldots \omega_M l_M S_M^k] \ (k = 0 \ldots A_0)\\
\mathbf{L}_{jk} &= \begin{bmatrix} \omega_1 l_1 S_1^k \sin j\omega t_1 \ldots \omega_M l_M S_M^k \sin j\omega t_M \\ \omega_1 l_1 S_1^k \cos j\omega t_1 \ldots \omega_M l_M S_M^k \cos j\omega t_M \end{bmatrix} \begin{matrix} (j = 1 \ldots p) \\ (k = 0 \ldots A_j) \end{matrix}\\
\mathbf{M}_{0k} &= [\omega_1 m_1 S_1^k \ldots \omega_M m_M S_M^k] \ (k = 0 \ldots B_0)\\
\mathbf{M}_{jk} &= \begin{bmatrix} \omega_1 m_1 S_1^k \sin j\omega t_1 \ldots \omega_M m_M S_M^k \sin j\omega t_M \\ \omega_1 m_1 S_1^k \cos j\omega t_1 \ldots \omega_M m_M S_M^k \cos j\omega t_M \end{bmatrix} \begin{matrix} (j = 1 \ldots q) \\ (k = 0 \ldots B_j) \end{matrix}\\
\mathbf{N}_{0k} &= [\omega_1 n_1 S_1^k \ldots \omega_M n_M S_M^k] \ (k = 0 \ldots C_0)\\
\mathbf{N}_{jk} &= \begin{bmatrix} \omega_1 n_1 S_1^k \sin j\omega t_1 \ldots \omega_M n_M S_M^k \sin j\omega t_M \\ \omega_1 n_1 S_1^k \cos j\omega t_1 \ldots \omega_M n_M S_M^k \cos j\omega t_M \end{bmatrix} \begin{matrix} (j = 1 \ldots r) \\ (k = 0 \ldots C_j) \end{matrix}
\end{aligned} \tag{31, 32}$$

VOL. 16 1959

All element of $\mathbf{S}$ are $\leqslant 1$ in absolute magnitude. Equation (9) is solved for

$$\mathbf{a} = [u_{00} \ldots u_{0A_0}\, u_{10}\, u^*_{10} \ldots u_{1A_1}\, u^*_{1A_1} \ldots u_{pA_p}\, u^*_{pA_p}\, v_{00} \ldots v^*_{qB_q}\, w_{00} \ldots w^*_{rC_r}] \quad (33)$$

The amplitudes of the jth harmonics are

$$\begin{aligned} U_j(z) &= [u_j{}^2(z) + u_j{}^{*2}(z)]^{1/2} \\ V_j(z) &= [v_j{}^2(z) + v_j{}^{*2}(z)]^{1/2} \qquad (j = 1 \ldots p) \\ W_j(z) &= [w_j{}^2(z) + w_j{}^{*2}(z)]^{1/2} \end{aligned} \quad (34)$$

and their phase angles are

$$\begin{aligned} \xi_j(z) &= -\tan^{-1} u_j(z)/u_j{}^*(z) \\ \eta_j(z) &= -\tan^{-1} v_j(z)/v_j{}^*(z) \qquad (j = 1 \ldots p) \\ \zeta_j(z) &= -\tan^{-1} w_j(z)/w_j{}^*(z) \end{aligned} \quad (35)$$

The choice of the period $2\pi/\omega$ would depend on the time interval for which observations are taken, and on a foreknowledge of the frequency components that are likely to be detected on the basis of the data. For example, if observations extend over only a few hours, time variations might be neglected and the analysis of Section 6 applied. With observations extending over a few days, the 24- and 12-hr components should be sought and possibly also the second harmonic: in this case, $2\pi/\omega$ would be taken as 24 hr with t in hours. From observations extending over a few months, detection of the lunar semidiurnal component might be considered: in this case, $2\pi/\omega$ would be taken as 28 days with t measured in days, and $j = 28$, 56 would be the diurnal and semidiurnal components, while $j = 58$ would be the lunar semicomponent (since this introduces components like $\sin 58\omega t$, which equals $\sin 4\pi\tau$, where τ is in lunar days).

The number of parameters involved in (33) is

$$[(A_0 + 1) + 2\sum_{j=1}^{p}(A_j + 1)] + [(B_0 + 1) + 2\sum_{j=1}^{q}(B_j + 1)] + [(C_0 + 1) + 2\sum_{j=1}^{r}(C_j + 1)]$$

If 24-, 12- and 8-hr components are sought in conjunction with a cubic variation in z for u, v, while w is taken to be independent of z and t, the number of parameters is 57.

8. Application of Theory

The results of Section 6, which deals with the case when all observations occur within a short interval of time, have been applied to thirty-three observations on meteor trails, made within the same hour of solar time. The observations were actually made over five consecutive days and the day-to-day variations at this particular hour were disregarded. Nearly half of the observations were between 85 and 89 km, the numbers in the height ranges 75–79, 80–84, 85–89, 90–94, 95–99 and 100–104 km being 3, 5, 14, 5, 3 and 3, respectively.

The variations with height of the EW and NS wind components were taken to be quadratic

$$u = u_0 + u_1 S + u_2 S^2$$
$$v = v_0 + v_1 S + v_2 S^2$$
$$S = (z - 96)/17$$

where z is the height of the reflection point in kilometers, and a constant vertical wind w_0 was assumed. The matrix $\mathbf{S}$ given by (26) was formed and $(\mathbf{SS}')^{-1}$ was

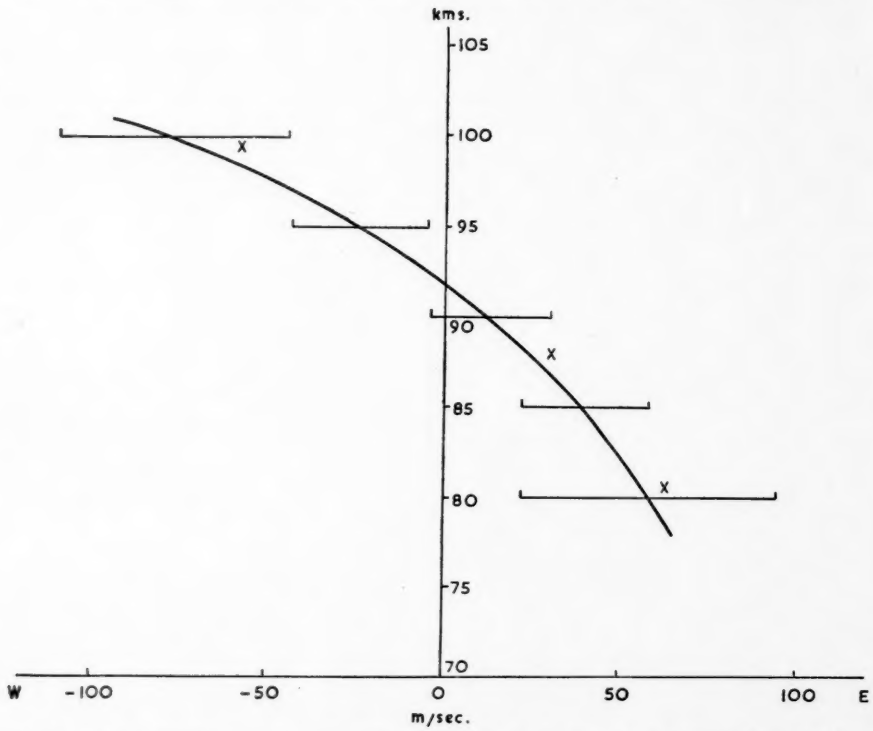


Fig. 1. EW wind component determined as a quadratic function of height and allowing for a constant vertical wind. The standard deviations are indicated. The crosses mark the values obtained by assuming the wind to be horizontal and constant in three adjacent height intervals.

evaluated. The seven parameters were obtained from (13) and their standard deviations from (A.6) as $u_0 = 35(\pm 20)$, $u_1 = 172(\pm 63)$, $u_2 = 75(\pm 85)$, $v_0 = 55(\pm 25)$ $v_1 = 27(\pm 73)$, $v_2 = 1(\pm 82)$, $w_0 = 8(\pm 8)$ m/sec. The EW and NS components are shown in Fig. 1 and 2. Their standard deviations are indicated by the bars, and are a minimum between 85 and 90 km where the observations have greatest density. Examination of the values just quoted shows that only u_0, u_1 and v_0 are significantly different from zero. The data could therefore be re-analysed by taking the EW component to vary linearly with height, the NS component as independent of height and the vertical component as zero. On the basis of the above analysis, the

VOL. 16 1959

gradient at the EW component is obtained as $7{\cdot}4(\pm 2{\cdot}2)$ m/sec per km at 91 km, the height of the centre of the distribution.

The three crosses on Figs. 1 and 2 indicate the values obtained by grouping the observations into the height intervals 75–84, 85–90 and 95–104 km and assuming that the wind is horizontal and independent of height within each of these height ranges. This is the case dealt with in Section 5. The crosses are placed at the

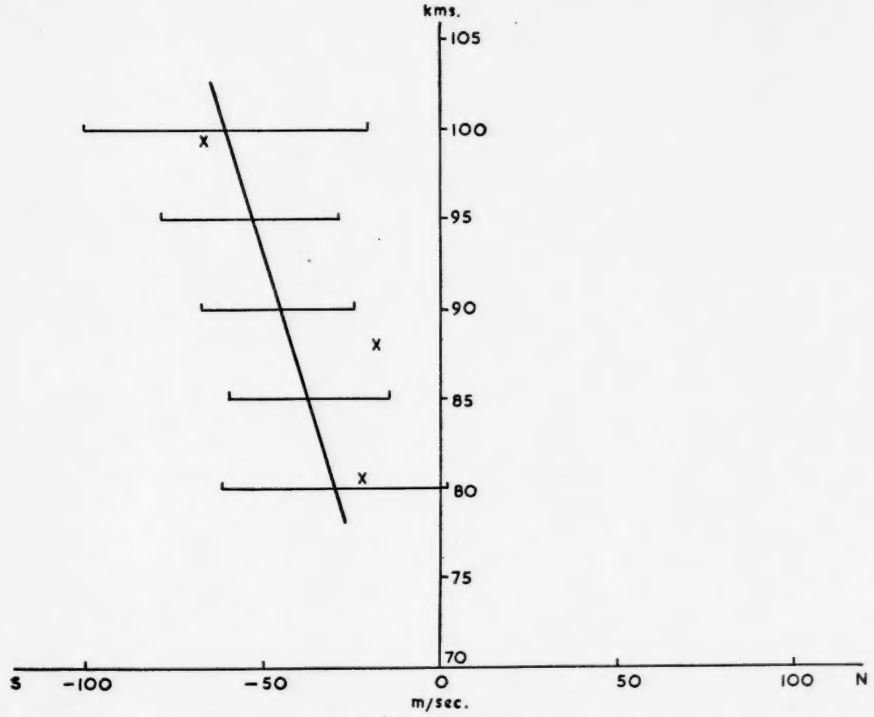


Fig. 2. NS wind component determined as a quadratic function of height allowing for a constant vertical wind. Symbols are the same as in Fig. 1.

average reflection height for each group. The difference between the two analyses is reasonably small, apart from the discrepancy of the central cross in Fig. 2, for which no explanation can be given.

In general, it would seem desirable to introduce as many parameters as practically possible into the analysis of a set of data so that unsuspected trends (e.g. harmonics in time variation) could be picked out and the accuracy of the other parameters improved. If many parameters were found to be redundant, the analysis could be repeated without them to give greater accuracy. When the same data is being analysed under various assumptions, the use of an electronic computer has obvious advantages and offers the possibility of analysing several thousands of observations simultaneously with a greater number of unknown parameters.

Acknowledgement—The calculations of Section 8 were carried out by Dr. W. G. Elford at the Department of Physics, University of Adelaide.

APPENDIX

1. *Distribution of* $\omega_i\varepsilon_i$

Consider a large number of hypothetical and apparently identical repetitions of the series of M observations. Reference to this "superpopulation" has already been implied in Section 2, where each ε_i was assumed to be distributed normally with standard deviation $\sigma\varepsilon_i$. Since $\omega_i \propto 1/\sigma\varepsilon_i$, $\omega_i\varepsilon_i$ $(i = 1, \ldots M)$ belong to the same normal distribution D with zero mean. The standard deviation of D is estimated by

$$\sigma^2\varepsilon_\omega = \sum_{i=1}^{M} (\omega_i\varepsilon_i)^2/\nu \tag{A.1}$$

where

$$\nu = M - (N_1 + N_2 + N_3) \tag{A.2}$$

is the number of independent observations. By (13) and (15)

$$[\omega_1\varepsilon_1 \ldots \omega_M\varepsilon_M] = \mathbf{aS} - [\omega_1 V_1 \ldots \omega_M V_m] \tag{A.3}$$

and hence $\sigma\varepsilon_\omega$ can be evaluated.

The standard deviation in the above estimate of $\sigma^2\varepsilon_\omega$ can be shown to be $(2/\nu)^{1/2}\,\sigma^2\varepsilon_\omega$ and the standard deviation in $\sigma\varepsilon_\omega$ can be shown to be $(4/\pi\nu)^{1/2}\,\sigma\varepsilon_\omega$. Hence to obtain $\sigma\varepsilon_\omega$ to 10 per cent accuracy ν needs to be in the region of a hundred. More precisely, statistical theory says that the $(1 - \alpha)$ 100 per cent confidence interval for $\sigma^2\varepsilon_\omega$ is

$$\left(\frac{\nu\sigma^2\varepsilon}{\chi_\nu^2{}_{,\alpha/2}} \quad \frac{\nu\sigma^2\varepsilon}{\chi_\nu^2{}_{,1-\alpha/2}}\right)$$

where $\chi_\nu^2{}_{,\alpha/2}$, $\chi_\nu^2{}_{,1-\alpha/2}$ are the values exceeded with probability $\alpha/2$, $1-\alpha/2$, respectively in the χ^2-distribution of ν degrees of freedom. For $\alpha = 0{\cdot}05$ and large $\nu(>100)$, the confidence interval is approximately $[1 \pm 1{\cdot}96(2/\nu)^{1/2}]\sigma^2\varepsilon_\omega$. These remarks apply to the standard deviations in the parameters a_{kn_k} as these will be seen to be proportional to $\sigma\varepsilon_\omega$ (equation (A.6)), and likewise to the standard deviations in u, v, w (equation (A.8)).

2. *Distribution of parameters,* a_{kn_k}

Let $\hat{\mathbf{a}}$ denote the values calculated from (13), and let $\mathbf{a}$ be the values corresponding to an arbitrary set of weighted residuals $\omega_1\varepsilon_1 \ldots \omega_M\varepsilon_M$
Then

$$\hat{\mathbf{a}} - \mathbf{a} = [\omega_1 V_1 \ldots \omega_M V_M]\,\mathbf{T} - \mathbf{a}$$

and hence by (6), (12) and (14)

$$\mathbf{a} = \hat{\mathbf{a}} + [\omega_1\varepsilon_1 \ldots \omega_M\varepsilon_M]\,\mathbf{T} \tag{A.4}$$

$\mathbf{a}$ is therefore normally distributed with mean value

$$\bar{\mathbf{a}} = \hat{\mathbf{a}} \tag{A.5}$$

Forming the matrix $(\bar{\mathbf{a}}' - \mathbf{a}')(\bar{\mathbf{a}} - \mathbf{a})$ and taking its mean we obtain a matrix whose diagonal elements are

$$\sigma^2 a_{kn_k} = \overline{(a_{kn_k} - \bar{a}_{kn_k})^2}$$

and whose non-diagonal elements are

$$\delta(a_{k_1n_{k_1}}, a_{k_2n_{k_2}}) = \overline{(a_{k_1n_{k_1}} - \bar{a}_{k_1n_{k_1}})(a_{k_2n_{k_2}} - \bar{a}_{k_2n_{k_2}})}$$

By (A.4) and (A.5), this matrix is

$$(\bar{\mathbf{a}}' - \mathbf{a}')(\bar{\mathbf{a}} - \mathbf{a}) = \mathbf{T}'\,\boldsymbol{\epsilon}_\omega'\boldsymbol{\epsilon}_\omega\,\mathbf{T}$$

where

$$\boldsymbol{\epsilon}_\omega = [\omega_1\varepsilon_1 \ldots \omega_M\varepsilon_M]$$

VOL. 16 1959

Hence on averaging and putting $\overline{\boldsymbol{\epsilon}_\omega' \boldsymbol{\epsilon}_\omega} = \mathbf{I}\sigma^2\varepsilon_\omega$, since $\omega_i\varepsilon_i$ are independently distributed,

$$\overline{(\bar{\mathbf{a}}' - \mathbf{a}')(\bar{\mathbf{a}} - \mathbf{a})} = \mathbf{T}'\mathbf{T}\sigma^2\varepsilon_\omega$$
$$= (\mathbf{S}\mathbf{S}')^{-1}\,\sigma^2\varepsilon_\omega \qquad \text{(A.6)}$$

from (14).

The diagonal elements of (A.6) give estimates for $\sigma^2 a_{kn_k}$ and the non-diagonal elements for $\delta(a_{k_1 n_{k_1}},\ a_{k_2 n_{k_2}})$, $\sigma^2\varepsilon_\omega$ having the value calculated from (A.1).*

To test whether the value of a_{kn_k} calculated from (13) is significantly different from zero, we compare it with σa_{kn_k}. If it is much greater than σa_{kn_k}, it can readily be accepted as significant, but if its value approaches the order of σa_{kn_k}, its significance becomes less certain. From statistical theory, the $(1 - \alpha)100$ per cent confidence interval for the mean of a_{kn_k} is

$$a_{kn_k} \pm \sigma a_{kn_k} t_{\infty,\alpha} \qquad \text{(A.7)}$$

where $t_{\infty,\alpha}$ is the value exceeded in absolute value with probability α in the normal distribution ($t_{\infty,\alpha} = 2{\cdot}6,\ 2{\cdot}0,\ 1{\cdot}2$ for $\alpha = 0{\cdot}01,\ 0{\cdot}05,\ 0{\cdot}25$). If the interval (A.7) does not contain the value zero, a_{kn_k} is significantly different from zero.

3. *Errors in wind components, magnitude and direction*

Write

$$u = \sum_{n_1} (\partial u/\partial a_{1n_1})\, a_{1n_1},$$

u being linear in the a_{1n_1}'s. Then $u - \bar{u} = \sum_{n_1} (\partial u/\partial a_{1n_1})(a_{1n_1} - \bar{a}_{1n_1})$ and

$$\sigma^2 u = (u - \bar{u})^2 = [\partial u/\partial a_{11} \ldots \partial u/\partial a_{1N_1}]\ \mathbf{S}_{11}[\partial u/\partial a_{11} \ldots \partial u/\partial a_{1N_1}]'\ \sigma^2\varepsilon_\omega$$

Likewise

$$\begin{aligned}
\sigma^2 v &= [\partial v/\partial a_{21} \ldots \partial v/\partial a_{2N_2}]\mathbf{S}_{22}[\partial v/\partial a_{21} \ldots \partial v/\partial a_{2N_2}]'\sigma^2\varepsilon_\omega \\
\sigma^2 w &= [\partial w/\partial a_{31} \ldots \partial w/\partial a_{3N_3}]\mathbf{S}_{33}[\partial w/\partial \alpha_{31} \ldots \partial w/\partial \alpha_{3N_3}]'\ \sigma^2\varepsilon_\omega \\
\delta(uv) &= [\partial u/\partial \alpha_{11} \ldots \partial u/\partial \alpha_{1N_1}]\mathbf{S}_{12}[\partial v/\partial \alpha_{21} \ldots \partial v/\partial \alpha_{2N_2}]'\ \sigma^2\varepsilon_\omega \qquad \text{(A.8)} \\
\delta(uw) &= [\partial u/\partial \alpha_{11} \ldots \partial u/\partial \alpha_{1N_1}]\mathbf{S}_{13}[\partial w/\partial \alpha_{31} \ldots \partial w/\partial \alpha_{3N_3}]'\ \sigma^2\varepsilon_\omega \\
\delta(vw) &= [\partial v/\partial \alpha_{21} \ldots \partial v/\partial \alpha_{2N_2}]\mathbf{S}_{23}[\partial w/\partial \alpha_{31} \ldots \sigma w/\partial a_{3N_3}]'\ \sigma^2\varepsilon_\omega
\end{aligned}$$

where

$$(\mathbf{S}\mathbf{S}')^{-1} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} & \mathbf{S}_{13} \\ \mathbf{S}_{12} & \mathbf{S}_{22} & \mathbf{S}_{23} \\ \mathbf{S}_{13} & \mathbf{S}_{23} & \mathbf{S}_{33} \end{bmatrix} \qquad \text{(A.9)}$$

For the wind magnitude $W = (u^2 + v^2 + w^2)^{1/2}$

$$W(W - \overline{W}) = u(u - \bar{u}) + v(v - \bar{v}) + w(w - \bar{w})$$

to the first order. Squaring and averaging gives

$$W^2\sigma^2 W = [u\ v\ w]\begin{bmatrix} \sigma^2 u & \delta(uv) & \delta(uw) \\ \delta(uv) & \sigma^2 v & \delta(vw) \\ \delta(uw) & \delta(vw) & \sigma^2 w \end{bmatrix}[u\ v\ w]' \qquad \text{(A.10)}$$

---

* If the observed value of times, heights and directions are randomly distributed, the diagonal elements of $\mathbf{S}\mathbf{S}'$ (which are $\sum_{i=1}^{M} (\omega_i l_i \partial u_i/\partial a_{kn_k})^2$ ($n_k = 1 \ldots N_k$), etc.) would be expected to be distributed with a mean value $\sim M$, for large M. Hence, the mean of the diagonal elements of $(\mathbf{S}\mathbf{S}')^{-1} \sim 1/M$; and, by (A.6), the standard deviations in a_{kn_k} (and also in u, v, w) $\sim 1/M^{1/2}$. On the other hand, the non-diagonal elements of $\mathbf{S}\mathbf{S}'$ would be expected to be distributed about zero, and so also would be those of $(\mathbf{S}\mathbf{S}')^{-1}$. The a_{kn_k}'s are then independently distributed.

Similarly, for the wind azimuth ϕ, where

$$\tan \phi = v/u,$$

$$\sec^2\phi\ (\phi - \bar{\phi}) = (v - \bar{v})/u - (u - \bar{u})v/u^2$$

$$\sec^4\phi\ \sigma^2\phi = \sigma^2 v/u^2 + \sigma^2 u(v^2/u^4) - \delta(uv)2v/u^3 \qquad \text{(A.11)}$$

i.e. $$\sigma^2\phi = [u^2\sigma^2 v + v^2\sigma^2 u - 2uv\delta(uv)]/(u^2 + v^2)^2$$

Similarly, for the wind elevation ψ, where

$$\tan \psi = w/(u^2 + v^2)^{1/2}$$

$$\sec^2\psi\ (\psi - \bar{\psi}) = \{(u^2 + v^2)(w - \bar{w}) - w[u(u - \bar{u}) + v(v - \bar{v})]\}/(u^2 + v^2)^{3/2}$$

$$\sec^4\psi\ \sigma^2\psi = \{(u^2 + v^2)^2\sigma^2 w + w^2[u^2\sigma^2 u + v^2\sigma^2 v + 2uv\delta(uv)] - 2w(u^2 + v^2)[u\delta(uw) + v\delta(vw)]\}/(u^2 + v^2)^3$$

$$\sigma^2\psi = \frac{1}{W^4(u^2 + v^2)}\,[uw\ vw\ -(u^2 + v^2)] \begin{bmatrix} \sigma^2 u & \delta(uv) & \delta(uw) \\ \delta(uv) & \sigma^2 v & \delta(vw) \\ \delta(uw) & \delta(vw) & \sigma^2 w \end{bmatrix} [uw\ vw\ -(u^2 + v^2)]' \qquad \text{(A.12)}$$

REFERENCES

BEYNON W. J. G. and BROWN G. M. (Editors) 1957 *IGY Instruction Manual, The Ionosphere* Part III, Chap. II. Pergamon Press, London

MANNING L. A., VILLARD O. G. and PETERSON A. M. 1950 *Proc. I.R.E.* **38**, 877.

VOL. 16 1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 357 to 359. Pergamon Press Ltd. Printed in Northern Ireland

The determination of the electron density in interplanetary space

J. M. KELSO
Space Technology Laboratories, Inc. Los Angeles 45, California

(*Received* 1 *March* 1959)

Abstract—A very high frequency (v.h.f.) pulse transmitted from the earth to a space vehicle at interplanetary distances will be subject to a delay exceeding that in free space by an amount proportional to the average electron density along the path. It is shown that this average electron density and the arc length of the ray path can be measured by transmitting a pulse to the vehicle on one frequency and transponding pulses from the vehicle on two frequencies. With sufficiently accurate measurements, the total electron content along the ray path through the ionosphere can be determined from such an experiment conducted from a high-altitude satellite.

PRESENT knowledge of the electron densities above the *F2*-layer maximum is very incomplete. As indicated by STOREY (1953) and later workers in the same field, observations of whistlers indicate that electron densities of the order of a few hundred to a thousand electrons per cubic centimetre are present out to distances of several earth's radii. Calculations based on observations of zodiacal light (see VAN DE HULST, 1956), indicate that electron densities of a similar order of magnitude may exist in the solar corona out to distances from the sun slightly less than that of the earth's orbit.

The existing capability of sending an instrumented payload into interplanetary space leads to the possibility of making direct observations relating to the electron density far from the earth. The present note shows that the measurement of the round-trip group path on at least two frequencies can establish the average electron density or, equally, the total electron density along the path and the arc length of the ray and, under suitable conditions, can yield data concerning the total electron content of the ionosphere.

At frequencies of the order of 50 to 100 Mc/s or greater, and for ionospheric electron densities, the index of refraction, μ, can be approximated very closely by the expression,

$$\mu^2 = 1 - \frac{K^2 N}{f^2} \tag{1}$$

where N is the electron density (electrons/m^3), f is the frequency (c/s), and $K^2 = 80{\cdot}5$ m^3-sec^{-2}.

Under these conditions, the group path may be written,

$$P' = \int_{\text{path}} \frac{ds}{\mu} \tag{2}$$

where ds is the element of arc length along the path, and where the integral is to be taken along the ray path.

Even for the electron densities existing at the maximum of the *F2*-layer, we have $K^2N/f^2 \ll 1$ at 50 Mc/s and above. Therefore, introducing the refractive index from equation (1) into equation (2), applying the binomial expansion, and neglecting squares and higher powers of (K^2N/f^2), we get, approximately,

$$P' = \int_{\text{path}} ds + \frac{K^2}{2f^2} \int_{\text{path}} N ds \tag{3}$$

The first integral clearly yields the arc length of the ray, which we shall denote as R, while the second integral gives the total number of electrons contained in a tube of 1 m^2 cross-section along the path.

To estimate the magnitudes involved, let us suppose that the electron density has the constant value of 10^9 electrons/m^3 (10^3/cm^3). Then the total number of electrons in a column having a 1 m^2 cross-section and a length of 10^6 km is equal to 10^{18} electrons. (It should be noted that this is only a factor of 2 or 3 greater than the number of electrons in a 1 m^2 vertical column through the ionosphere near midday as observed by EVANS (1957) using the Faraday rotation of pulses reflected from the moon.) Assuming a straight path, and using a frequency of 100 Mc/s, we see from equation (3) that the group path exceeds the arc length, R, by the amount of 4·0 km per million kilometres of path. Therefore, over a distance of 10^8 km the divergence between P' and R amounts to 400 km.

As a method for carrying out the experiment, let a pulse at frequency f_1 be transmitted from the earth to the vehicle. When the pulse is received in the vehicle, let two pulses at frequencies f_1 and f_2 be transmitted back to earth. Let the one-way group paths be $P'(f_1)$ and $P'(f_2)$. Then at the ground, the received pulse at frequency f_1 will have travelled an apparent path equal to $2P'(f_1)$, while the pulse at frequency f_2 will have travelled an apparent distance $P'(f_1) + P'(f_2)$. Since the distance R will be the same in both cases, we can use equation (3) to obtain,

$$2P'(f_1) - [P'(f_1) + P'(f_2)] = \frac{K^2}{2}\left(\frac{1}{f_1^2} - \frac{1}{f_2^2}\right)\int N ds \tag{4}$$

The left-hand side represents the difference in the group paths and is measurable. On the right-hand side, only the integral is unknown and can therefore be calculated from the experiment. It is seen that this calculation requires only the difference between the group paths. If also, we determine the total value of one of the group paths, say $2P'(f_1)$, then the value of R can be computed.

Thus we see that if we make measurements over a distance of 10^8 km using, say, frequencies of 100 and 400 Mc/s, and are able to measure the difference in group paths to within 20 to 40 km then an average electron density as low as 50–100 electrons/cm^3 will produce a measurable difference.

It might also be noted that use of a frequency as low as 50 Mc/s would produce a difference $P' - R$ of about 5 km on a one-way trip through the ionosphere at vertical incidence near summer noontime. Consequently, a group-path experiment might well be used with a high-altitude satellite or other space vehicle to obtain

VOL. 16 1959

information concerning the variations in the total electron content of the ionosphere.* This experiment would require somewhat greater accuracy in measurement of group-path differences than has been assumed for the interplanetary experiment.

The preceding discussion is intended only to indicate the general manner in which such an experiment can be carried out. The specific details of the experiment cannot be developed fully until a particular space vehicle system is established, since the experimental details must be adjusted to conform to the characteristics of the space vehicle, especially as concerns the available transmitter power. To conserve band-width, and hence transmitter power, detection schemes using correlation techniques might be employed. Other modifications, such as the introduction of known delays before transponding from the vehicle, will suggest themselves.

The above discussion has shown that a multifrequency group-path experiment between the ground and a space vehicle would make it possible to determine the average electron density in interplanetary space to within values of the order of 100 electrons/cm^3, to determine the actual arc length of the ray between the ground and the vehicle, and, under suitable circumstances, to obtain information concerning the variations in the total electron density in the ionosphere.

L.
6
59

References

Evans J. V. 1957 *J. Atmosph. Terr. Phys.* **11**, 259.

Storey L. R. O. 1953 *Phil. Trans.* **246**, 113.

Van de Hulst H. C. 1956 *Vistas in Astronomy* Vol. II, p. 998. Pergamon Press, London.

* Suggestions for an ionospheric group-path experiment, apparently similar to that proposed here, were made independently by S. Weisbrod and by G. H. Millman at the Fall 1958 URSI meeting in State College, Pennsylvania. This information was obtained from brief comments contained in a rough draft of a report by W. C. Hoffman concerning a symposium on the use of space vehicles, held at the URSI meeting. Publication of Dr. Hoffman's report is anticipated.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 360 to 366. Pergamon Press Ltd. Printed in Northern Ireland

The effects of a solar eclipse on a stratified ionosphere

J. A. GLEDHILL
Department of Physics, Rhodes University, Grahamstown, South Africa

(*Received* 9 *March* 1959)

Abstract—The effects of an eclipse on a simple three-layer ionosphere, in which the recombination coefficient decreases with height, have been computed. The results are displayed in the form of a contour map and show relatively large tilts of the isoelectronic surfaces near the layer maxima. The development of a "valley" between the *F1*- and *F2*-layers after the maximum of the eclipse is of particular interest. It is pointed out that such a valley may cause considerable errors in the interpretation of the behaviour of the *F2*-layer.

IT has been suggested by MUNRO and HEISLER (1958) that curvature of the isoelectronic surfaces may cause serious errors in the interpretation of ionosphere records taken during eclipses. MINNIS (1958) has put forward the view that such errors, while undoubtedly present, are not significant. His grounds for this belief are open to criticism, however, and records taken in South Africa during the eclipse of 25 December 1954 show effects which may be due to oblique reflections of the kind suggested by MUNRO and HEISLER (GLEDHILL, 1959). It is desirable to have some guide as to the likely effects which may be produced in the $h'f$ records.

In order to study these phenomena, a typical electron distribution was chosen and the effects of an eclipse on it have been computed, under certain simplifying assumptions.

THE IONOSPHERE MODEL

The electron distribution selected was based on that derived from a control day record at Grahamstown, at about the time of day at which the eclipse took place, by SZENDREI and MCELHINNY (1956, Fig. 1). It was simplified by eliminating

Table 1. Data for the ionosphere model

Height (km)	Electron density (cm^{-3})	α (cm^3/sec)
300	$49{\cdot}0 \times 10^4$	$2{\cdot}5 \times 10^{-11}$
280	48·0	3
260	44·0	5
240	38·0	15
220	24·0	40
200	23·0	150
180	22·5	500
160	19·5	500
140	14·0	700
130	10·3	1000
120	10·0	1000
110	7·0	1000
100	1·1	1000

VOL. 16 1959

the *E2*-layer and making the maximum electron density in the *E*-layer exactly 10^5 cm^{-3}. Since SZENDREI and MCELHINNY's records were not spread in height sufficiently to allow them to observe the shape of the *E*-layer, it was assumed to be a Chapman layer with a scale height of 10 km and the maximum at 120 km.

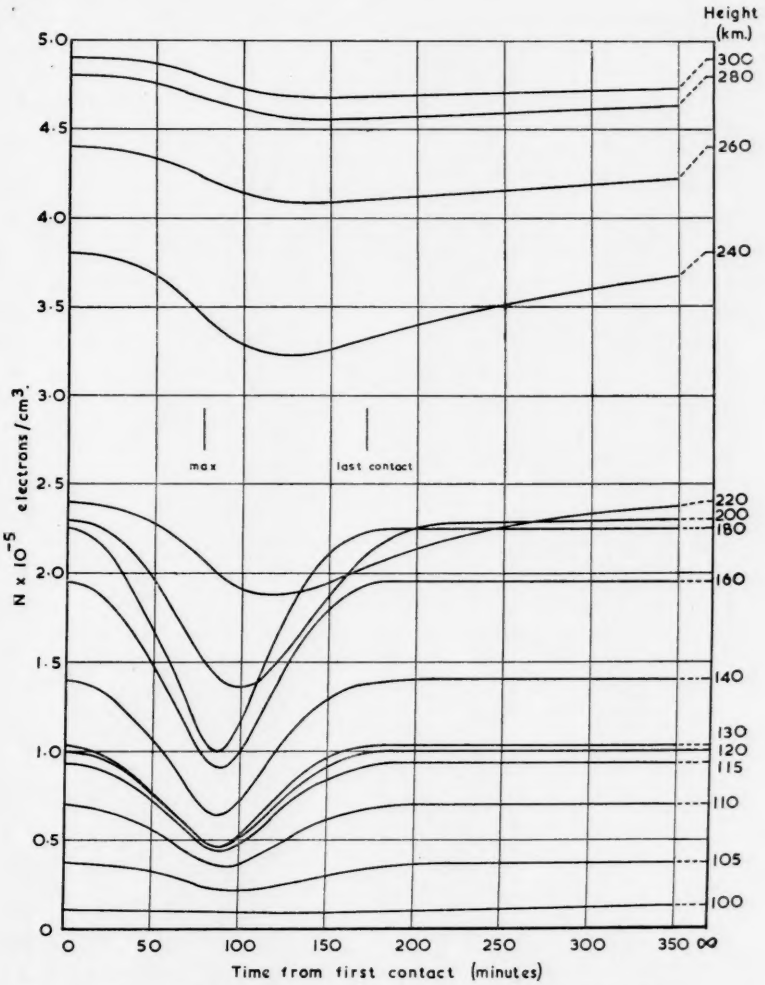


Fig. 1. Curves of electron density at selected constant heights.

This was joined smoothly on to the bottom of the *F1*-layer with no valley between. Recombination was assumed to be effective at all heights, rather than attachment, and the recombination coefficients used were those found by SZENDREI and MCELHINNY (1956, Table 2) modified somewhat in the transition regions between the layers so as to lie on a reasonably smooth curve. The figures used are listed in Table 1 and form the curve for $t = 0$ in Fig. 2.

The electron density at any fixed height was assumed to follow the equation

$$\frac{dN}{dt} = fq - \alpha N^2 \tag{1}$$

where N is the electron density at time t, when f is the fraction of the sun's area not covered by the moon, q is the rate of electron production in the absence of an eclipse and α is the recombination coefficient. q was calculated on the assumption

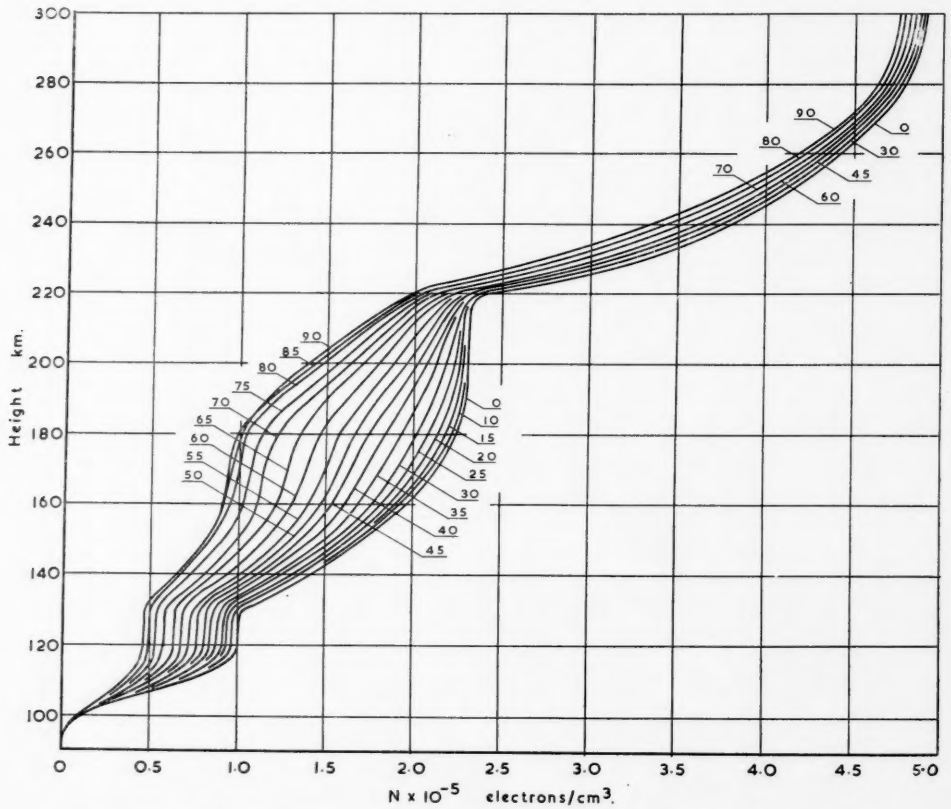


Fig. 2. N–h curves at given times (in minutes) after first contact: (a) during the first part, (b) during the latter part of the eclipse (p. 363).

that the ionosphere had reached a steady state before the onset of the eclipse, i.e. $dN/dt = 0$; it was assumed to remain constant throughout the eclipse, which is equivalent to neglecting the rotation of the earth. The values used for f were those computed for the annular eclipse at Grahamstown on 25 December 1954.

RESULTS

Equation (1) was solved numerically step by step throughout the eclipse at selected heights. The resulting curves are shown in Fig. 1. The time lag between

VOL. 16 1959

the maximum of the eclipse and the minimum of electron density is seen to increase with height as α decreases, as expected. This is, in fact, a more quantitative demonstration of the situation illustrated by RATCLIFFE (1956, Fig. 2) in his exposition of the possible origin of the *F1*½-layer. The implications of the crossing of the curves for 180, 200 and 220 km are commented upon later. The time lag is also observed to increase with decreasing height below 120 km. This is a consequence of the choice of a constant α in this region.

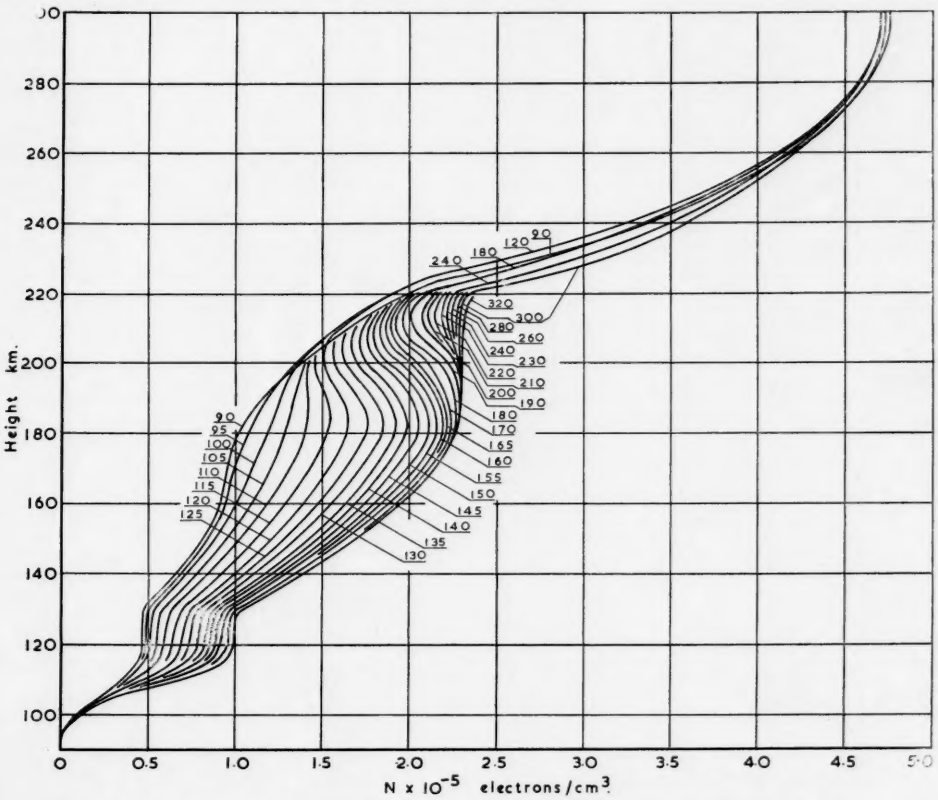


Fig. 2. *N*–*h* curves at given times (in minutes) after first contact: (a) during the first part (p. 362), (b) during the latter part of the eclipse.

From these curves, electron density profiles of the ionosphere at any desired time may be found. The family showing the densities during the first 90 min of the eclipse is shown in Fig. 2(a) and the remainder in Fig. 2(b). The latter shows clearly how the lower part of the *F1*-layer recovers while its upper part is still decreasing in density, corresponding to the crossing-over of the curves in Fig. 1. This is, of course, a consequence of the choice of a very small density gradient between 180 and 220 km in the original model. The lower maximum is not the *F1*½-layer, however, since it is the only maximum between the *E*- and *F2*-layers.

The "valley" between the *F1*- and *F2*-layers would produce serious errors if *h'f* records during this part of the eclipse were converted to true heights by any

of the normal methods. On the usual assumption of a constant density above the *F1* maximum, all "true" heights in the *F2*-layer would be too low. As a result curves of ionization at fixed heights would show too quick a recovery after the eclipse; they may even show a maximum during the later part of the eclipse and regain normal values only some time afterwards when the valley had disappeared. This is what happened in the curves derived by SZENDREI and MCELHINNY (1956, Figs. 3, 6 and 7), and in two of the eclipses discussed by THOMAS and ROBBINS (1956). The presence of such a valley is a possible explanation. The results obtained by MINNIS (1956) for the Khartoum eclipse of 25 February 1952 show a much quicker rate of recovery than theory predicts at heights of 250, 300 and 350 km, which may be due to a similar cause.

It is quite possible that a corresponding valley may develop between the *E*- and *F1*-layers. It did not do so in the present example because of the particular choice of values in Table 1, but a valley could have been produced by choosing a recombination coefficient which decreased more rapidly with height just above the *E*-layer. Thus ionization curves in the *F1*-region may also be affected by errors of this nature. It is evident that deductions of recombination and attachment coefficients, and also of movements in the upper ionospheric layers, should be regarded with caution until it has been shown that the measurements were not affected by such a phenomenon. Errors due to this cause may well be as important as possible ones due to tilted layers.

VOL. 16 1959

MAP OF IONOSPHERIC CHANGES DURING ECLIPSE

If now several further assumptions are made, it is possible to construct from the given curves a map showing the ionosphere at points along the central line of the eclipse. We assume that the earth's curvature can be neglected and that the undisturbed ionosphere has the same form as our model at all points before the eclipse starts; surfaces of constant electron density would thus be planes parallel to the surface of the earth.

The curves in Figs. 2(a) and 2(b) now represent the changes with time at each point along the path of the shadow. Assuming a constant typical velocity of 40 km/min west to east for the shadow, it will be evident that at a place *B*, 40 km to the west of a given spot *A* on the central line of the eclipse, the state of the ionosphere at a given instant is that which will apply at *A* in one minute's time; while at a place *C*, 40 km to the east of *A*, the eclipse is 1 min less advanced than it is at *A*. Thus by converting time to distance at the rate of 40 km/min, a contour map may be deduced from the curves of Figs. 1 and 2, showing isoelectronic lines throughout the eclipse.

Since the relation of the vertical and horizontal scales is fixed by the shadow velocity, the map is much longer than it is high. With the scale of 1 in. to 20 km chosen it is more than 30 ft long. Fig. 3(a) shows a simplified version of its central portion and Fig. 3(b) the part during which the valley develops between the *F1*- and *F2*-layers.

These figures illustrate clearly the tilts produced in the *E*-, *F1*- and *F2*-layers near their electron density maxima. The actual slopes of the isoelectronic lines are very sensitive to the gradients originally assumed, but those shown seem to be

reasonable. The "inverted saucer" shape of the *E*-region overhead at mid-eclipse is evident, but one cannot say whether oblique reflections would be received from the "shoulders" without careful ray tracing. It does appear, from simple geometrical optics, that reflections could be returned from angles as large as 10° to the vertical. The rather abrupt changes in critical frequency which occur in actual eclipse records would correspond to even greater tilts of the isoelectronic

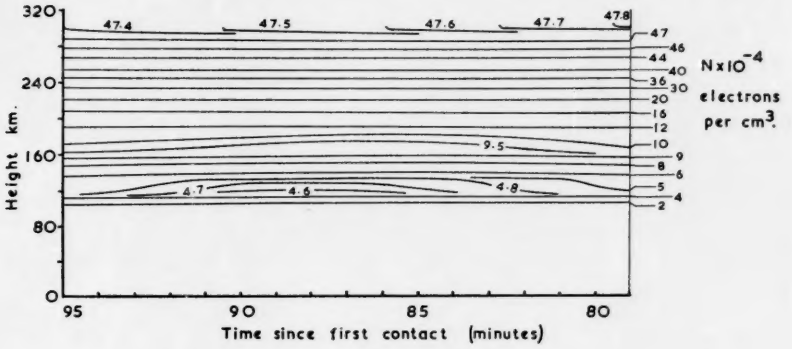


Fig. 3a.

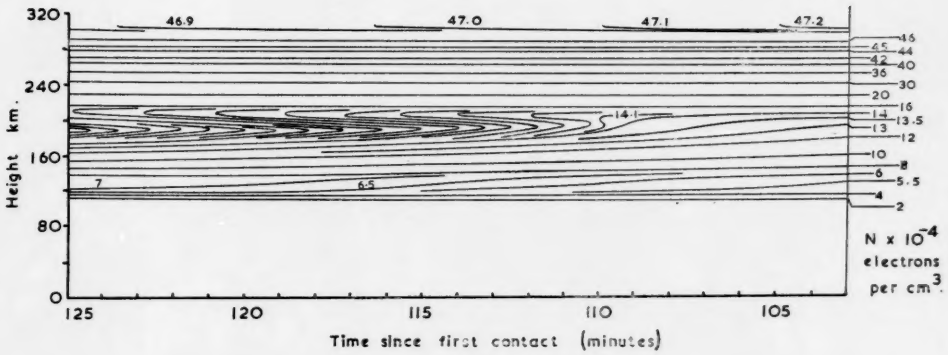


Fig. 3b.

Fig. 3. Isoelectronic contours along the central line of the eclipse: (a) at maximum eclipse, (b) showing the later development of the "valley" between the *F1*- and *F2*-regions.

curves. Reflections from the "end" of the post-eclipse *F1* maximum (Fig. 3(b)) are possible at all angles, though they would decrease in intensity with distance from it.

It may appear that the neglect of the earth's curvature would invalidate any quantitative ray-tracing on the map, and it is indeed quite inaccurate over large distances. Echoes are returned to a given point on the ground, however, only for relatively small angles of incidence, up to about 15°, and for such reflections the correction due to curvature of the earth would be very small. The expected forms of *h'f* records during the eclipse can thus be deduced by ray-tracing on the map. By comparing these with the records which would result from a plane stratified ionosphere, and by scaling them in the normal way to see what true

height distributions would result, the effects of tilts in the layers and of the development of the valley can be studied.

REFERENCES

GLEDHILL J. A.	1959	*J. Atmosph. Terr. Phys.* **16,** 367.
MINNIS C. M.	1956	*Solar Eclipses and the Ionosphere* p. 81. Pergamon Press, London.
MINNIS C. M.	1958	*J. Atmosph. Terr. Phys.* **12,** 272.
MUNRO G. H. and HEISLER L. H.	1958	*J. Atmosph. Terr. Phys.* **12,** 57.
RATCLIFFE J. A.	1956	*Solar Eclipses and the Ionosphere* p. 1. Pergamon Press, London.
SZENDREI M. E. and MCELHINNY M. W.	1956	*J. Atmosph. Terr. Phys.* **9,** 118.
THOMAS J. O. and ROBBINS A. R.	1956	*Solar Eclipses and the Ionosphere* p. 94, Pergamon Press, London.

VOL.
16
1959

Journal of Atmospheric and Terrestrial Physics 1959, Vol. 16, pp. 367 to 375. Pergamon Press Ltd. Printed in Northern Ireland

The behaviour of the ionosphere over Cape Town and Johannesburg during the annular solar eclipse of 25 December 1954

J. A. GLEDHILL
Department of Physics, Rhodes University, Grahamstown, South Africa

(*Received* 9 *March* 1959)

Abstract—Data are presented showing the variation of the critical frequencies for the ordinary ray in the *E*-, *F1*- and *F2*-regions at Cape Town and Johannesburg during the eclipse of 25 December 1954. The maximum obscuration was 79 per cent at Cape Town and 68 per cent at Johannesburg. It was not possible to deduce a solar model which would account satisfactorily for the data from both stations and those from Grahamstown, published earlier. Reasons are given for believing that oblique reflections affected the records and it is pointed out that the plot of f_0E vs. f_0F1, used by MINNIS to justify neglect of these reflections, is in fact not linear throughout an eclipse. There is some evidence for a high value of the recombination coefficient, 4×10^{-8} cm^3/sec, in the *E*-layer.

INTRODUCTION

THE behaviour of the ionosphere over Grahamstown during the eclipse of 25 December 1954 has been discussed in detail by SZENDREI and MCELHINNY (1956a, b, c). Grahamstown lay in the centre of the belt of annularity for ionospheric heights and 85·8 per cent of the area of the solar disk was covered during the annular phase.

During the same eclipse, ionosondes at Cape Town and Johannesburg, belonging to the National Telecommunications Research Laboratory of the Council for Scientific and Industrial Research, were operated at 1 min intervals. These instruments are of the fast-sweeping panoramic type (WADLEY, 1949) completing each sweep in 7 sec and recording on 16 mm film. Cape Town lay about 200 km south of the annularity belt, while Johannesburg lay about 600 km north of it. The circumstances of the eclipse at both places are given in Table 1.

Table 1. Circumstances of eclipse at Cape Town and Johannesburg
(South African Standard Time = UT + 2 hr)

Location	Height (km)	First contact	Max.	Last contact	Area covered at max. (%)
Cape Town					
18° 19′ E	100	0647	0759	0922	79·4
34° 09′ S	200	0646	0757	0921	78·2
	300	0644	0756	0919	77·0
Johannesburg					
28° 05′ E	100	0640	0756	0928	67·9
26° 10′ S	200	0638	0753	0927	68·9
	300	0636	0752	0925	70·1

RESULTS

In the interpretation of ionospheric eclipse records it is usually found that the variations of critical frequency do not follow a smooth curve. These irregularities are frequently ascribed to non-uniformities in the distribution of the sources of ionizing radiation on the solar disk. It is therefore important to assess the limits of error of the critical frequencies so as to be able to decide whether a particular value does or does not deviate significantly from that predicted by the hypothesis under test. In scaling the records from Cape Town and Johannesburg, therefore,

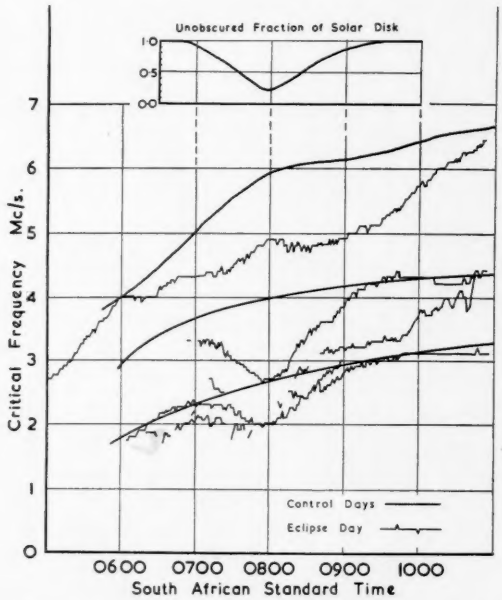


Fig. 1. Critical frequencies observed at Cape Town.

each critical frequency was noted together with its maximum and minimum possible values.

Many of the records from both stations show extra cusps. The region between the *E*- and *F1*-layers at Cape Town in particular appears to have been rather complex. It is suggested later in this paper that these extra cusps are probably due to oblique reflections, produced in the way discussed by MUNRO and HEISLER (1958). During the scaling of the records the frequencies of all such cusps were read as if they were normal penetration frequencies. Fig. 1 shows the frequencies of all cusps on the Cape Town records and Fig. 2 those for Johannesburg. The control day curves also shown are drawn through the means of the hourly values for the ten days centred on the eclipse day, published by the NTRL (NTRL, 1955).

The *E*- and *F1*-layers at both places show definite eclipse effects and the complexities are shown by the extra "layers". It is obvious, from these graphs, which cusps correspond to the normal *E*- and *F1*-layers, except for the period between 0840 and 0900 hours at Cape Town, where there is confusion about the

VOL. 16 1959

E-layer. The *F2* critical frequencies are, as has often been the case, inconclusive. The Cape Town figures fall considerably below the control values, but show a maximum in the middle of the eclipse. At Johannesburg, electron production appears to be faster during parts of the eclipse than before or after it.

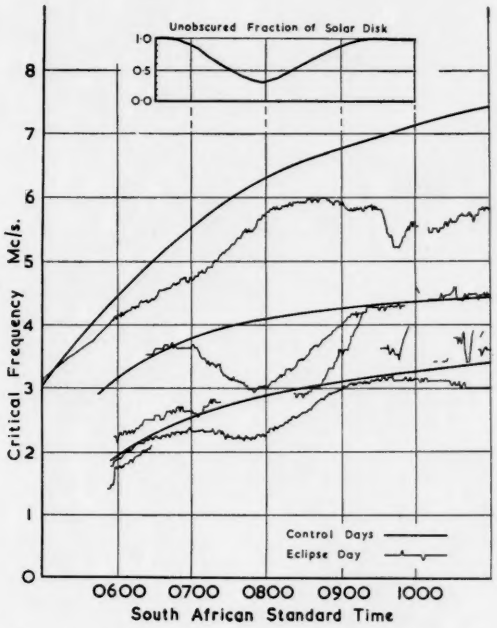


Fig. 2. Critical frequencies observed at Johannesburg.

Discussion

Fig. 3 compares the experimental results for the maximum electron density in the *E*-layer at Cape Town with theoretical curves calculated in the usual way from the control data for various values of the recombination coefficient assuming uniform radiation from the solar disk. The observed electron densities are represented by vertical lines so that the true value is definitely known to lie between their upper and lower ends. None of the theoretical curves fits the observations and obviously no other one based on the assumption of a uniform disk would do so. The corresponding Johannesburg data are shown in Fig. 4. The records taken at Johannesburg were of unusually good quality, and it was often possible to establish critical frequencies to within 0·01 Mc/s either way. Again no curve based on a uniform disk could possibly fit the experimental data. Figs. 5 and 6 compare the *F1*-layers at the two places with similar theoretical curves and lead to similar conclusions.

An attempt was made to use the Cape Town and Johannesburg data to decide the ambiguity in Szendrei and McElhinny's solar model (1956b), in which the bright sources may actually lie in either the northern or southern solar hemispheres and still be consistent with the observations at Grahamstown. The times of the

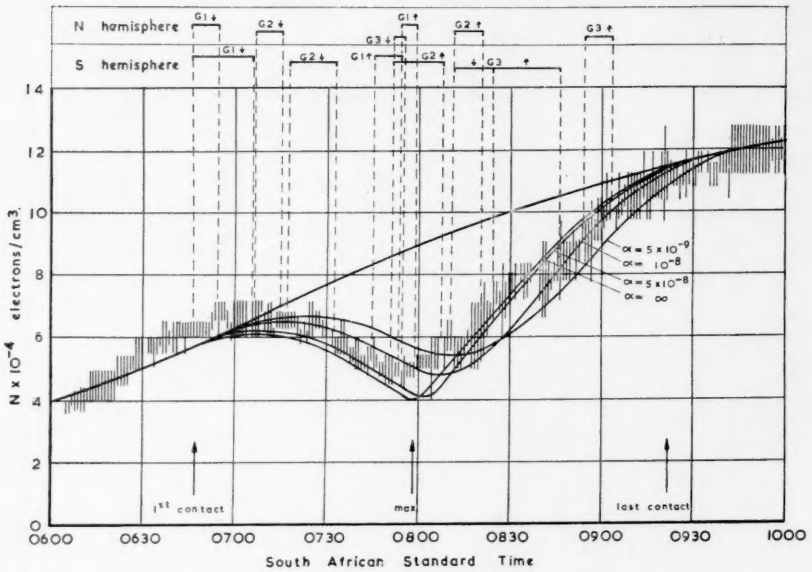


Fig. 3. Variation of electron density in the E-layer at Cape Town and comparison with SZENDREI and MCELHINNY's solar model.

VOL. 16 1959

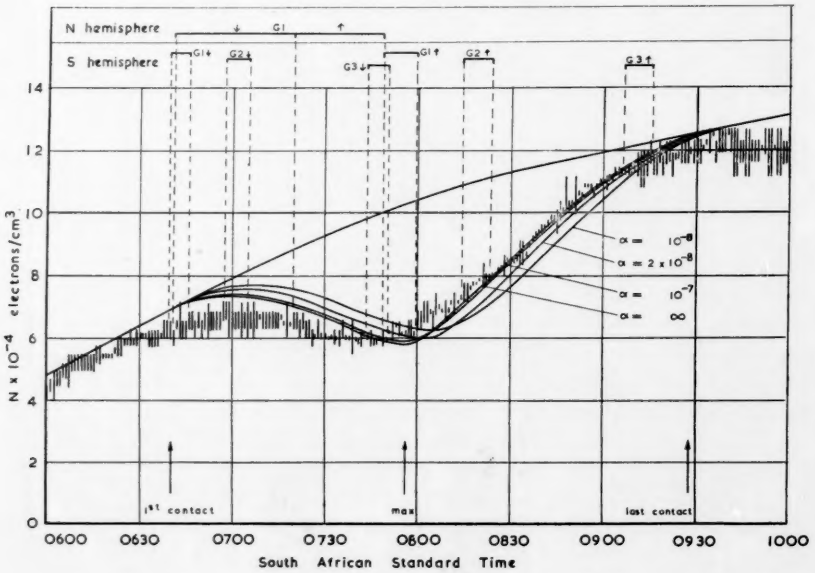


Fig. 4. Variation of electron density in the E-layer at Johannesburg and comparison with SZENDREI and MCELHINNY's solar model.

occultation and reappearance of all three sources are shown in Figs. 3 and 4, for both their northern and southern positions. There is some agreement with sources-*G1* and -*G2* in their northern positions but it is not nearly as conclusive as one would have liked. Solar models derived to fit the Cape Town data alone,

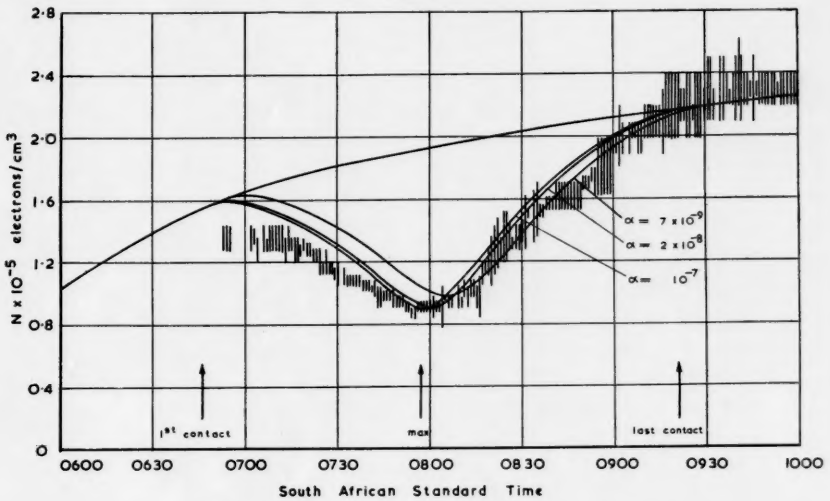


Fig. 5. Variation of electron density in the *F1*-layer at Cape Town.

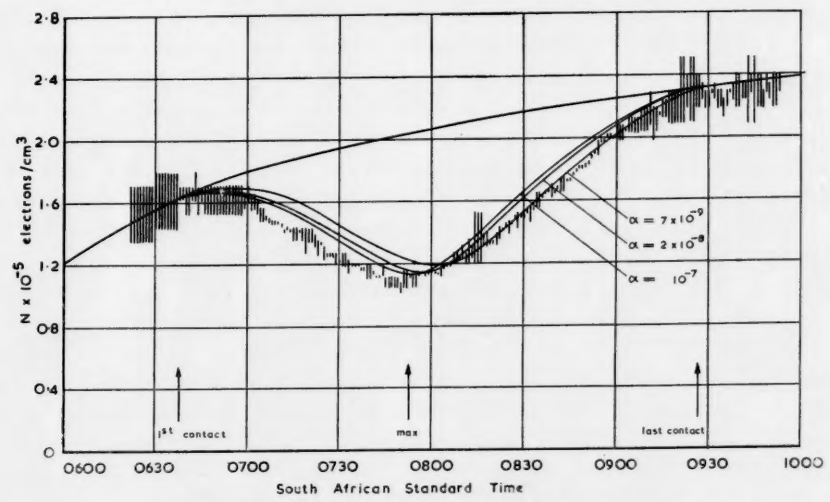


Fig. 6. Variation of electron density in the *F1*-layer at Johannesburg.

and the Johannesburg data alone, show no resemblance to each other or to that deduced by SZENDREI and MCELHINNY from the Grahamstown data. An attempt was also made to derive a distribution which would fit all three sets of data, by drawing arcs on a map of the sun, corresponding to the position of the moon's limb when sharp changes of gradient occurred on the electron density graphs.

In the ideal case three disappearance arcs should cross at a bright spot and three reappearance arcs should intersect at the same place. In fact there was only one area where three disappearance arcs did intersect, and this did not coincide with any reappearance arc nor with any observed feature on the sun's disk.

It seems almost certain that the observations at this eclipse are subject to errors due to curvature of the ionospheric layers, of the type suggested by MUNRO and HEISLER (1958). Many of the records show extra cusps of the expected type, and the extra "stratification" shown in Figs. 1 and 2 corresponds with that shown in their Fig. 6. The peculiar cessation in the rate of increase of the *E*-layer density at Cape Town between 0840 and 0850 hours appears to be associated with the appearance of the two extra cusps between *E* and *F1*, and the corresponding flat portion of the *F1*-graph for Cape Town may well be connected with the same phenomenon.

MINNIS (1958) has recently considered the criticism by MUNRO and HEISLER (1958), that solar models derived from ionospheric eclipse data may be in error because of the effects of distortion of the layers. He has concluded that such distortions, while theoretically present, produce only small errors in the interpretation. His main ground for this belief is the experimental observation that the ratio of the critical frequencies of the *E*- and *F1*-layers is constant throughout the course of each of the four eclipses which he considers. One would expect the deviations from vertical reflection to be greater in a thick layer such as the *F*1 than in a thin one such as the *E*, and thus would anticipate a breakdown of the linear relationship during the eclipse if such deviations were important.

VOL. 16 1959

A little consideration, however, shows that this linear relationship cannot hold for all eclipses, for the following reason. On normal days, when there is no eclipse, the critical frequencies of the *E*- and *F1*-layers are linearly related, but they fall on a straight line which does not pass through the origin. This is illustrated in Fig. 7 by the control day data for all three stations observing the present eclipse, and has also been shown by MINNIS (1956, 1958). During a normal day the point representing conditions in the two layers moves up this line during the morning, as the solar zenith distance decreases, and down again in the afternoon. On an eclipse day, if MINNIS's contention is true, this point must leave the normal line as soon as the eclipse commences and move down a new straight line which passes through the origin (since the ratio f_0F1/f_0E must now remain constant). It must remain on this new line throughout the eclipse, and therefore must return to the control line, at the end of the eclipse, at the point where it left it. Now in general the solar zenith distance will have changed considerably during the eclipse, and therefore the point will have returned to a part of the control line which does not correspond to the solar zenith distance at that time. One is left to wonder how it eventually catches up again so that the variation during the rest of the day may be normal.

On plotting the data from the stations observing the present eclipse (Fig. 7), it is at once seen that the ratio of critical frequencies does not remain constant at Johannesburg and Grahamstown (cf. MCELHINNY, 1958). Before the eclipse the points lie on the normal curve. When the eclipse gets under way they deviate from it, and f_0F1 is seen to fall much more rapidly than f_0E. After the eclipse

maximum their ratio does become constant, and they eventually regain the normal line, at the correct place for the solar zenith distance at the end of the eclipse, by following a straight line which now passes through the origin. The lines through the eclipse points in Fig. 7 have in fact been drawn so as to pass through the origin. Although the data for Cape Town do not show this behaviour during the first part of the eclipse, this seems to be due to the fact that Cape Town lies some 800 km west of Grahamstown. Thus the *F1*-layer was not in evidence at all at

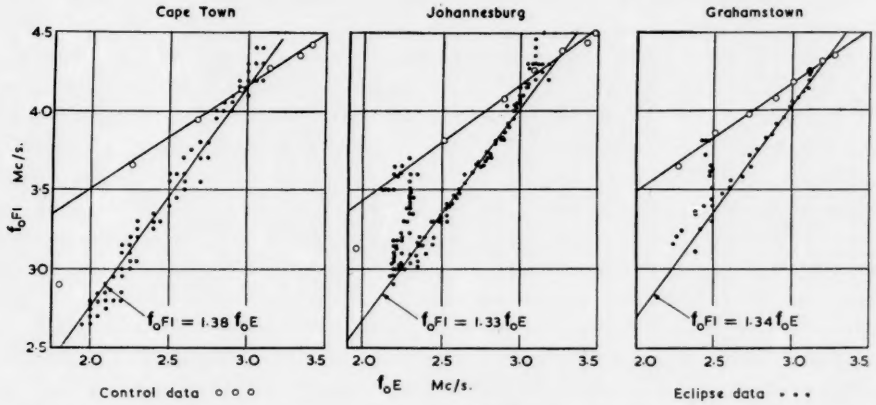


Fig. 7. Relations between critical frequencies in the *E*- and *F1*-layers. (Control-day points at each hour from 0600 to 1200 for Cape Town and Johannesburg and each half hour from 0630 to 1000 for Grahamstown).

Cape Town when the eclipse started, and was first seen as a change in gradient at 0702 hours, 16 min after first contact at 200 km. At Johannesburg, on the other hand, the *F1*-layer appeared at 0622 hours, and showed a definite cusp by the time of first contact at 0638 hours. At Grahamstown the circumstances were intermediate, but three approximate values for f_0F1 were readable before the eclipse commenced and these lie on the normal line. The eclipse at Cape Town thus started too early for normal *F1*-conditions ever to be established.

It appears that the reason why MINNIS's data do not show this effect is that all the eclipses, which he considered, took place near to local noon, so that there was not much difference in the solar zenith distance before and after them.

It is further worthy of note that, although MINNIS states that there is in no case a significant departure from the linear relation, some of the points shown in his Fig. 5 do lie well off the straight lines, e.g. the Khartoum value at $f_0E = 2{\cdot}37$ Mc/s, $f_0F1 = 2{\cdot}84$ Mc/s. When converted to electron densities, this and its nearest point on the straight line become $6{\cdot}95$ and $6{\cdot}55 \times 10^4$ cm^{-3}, respectively, a difference which is of the same order of magnitude as those which are interpreted as due to bright sources on the sun. There are several other points which lie comparably off the graph. In view of this, and the above demonstration of the failure of the proportionality between f_0E and f_0F1, it is felt that the disturbances suggested by MUNRO and HEISLER may well be of importance in the interpretation of eclipse effects, and the present data are currently being studied from that point of view.

The phenomenon shown in Fig. 7 is of considerable significance for the discussion of eclipse effects. The rapid fall in f_0F1 at the beginning of the eclipse seems to imply some very efficient mechanism for the removal of electrons. Its direct proportionality to f_0E during the second part of the eclipse requires the ratio of recombination coefficients in the two layers, and also the ratio of the maximum rates of electron production, to be constant and equal to each other, at least according to simple theories. It is desirable to study similar graphs for eclipses occurring at different times of day to confirm the generality of the relationships. The writer hopes to publish a further paper on this shortly.

MCELHINNY (1959) has considered the data presented in this paper for the *E*-layer. He has shown that the observations at all three places are reasonably consistent with a solar model in which 10 per cent of the radiation comes from the neighbourhood of a sunspot and 15 per cent from outside the east and west limbs of the sun. The fit is acceptable during the central portion of the eclipse, but not very good after 0800 hours for the Johannesburg and Cape Town figures and between 0730 and 0750 hours for the Grahamstown ones. The main attractive feature of this type of explanation is that it accounts well for the early and flat minimum in the experimental graphs. The recombination coefficient deduced, 4×10^{-8} cm^3/sec, is considerably higher than is usually obtained by other methods. BONNET *et al.* (1957) have recently, however, considered data from several stations observing the *E*-layer during the eclipse of 25 February 1952, assuming the presence of radiation from outside the visible solar disk, and have deduced the same value, 4×10^{-8} cm^3/sec, for the recombination coefficient.

VOL. 16 1959

The data for the *F1* critical frequencies in Figs. 5 and 6 do not support a similar explanation for the *F1*-layer. Both sets of observations, indeed, lie below the theoretical curve for $\alpha = \infty$ during the first part of the eclipse, and the same is true for the *F1*-layer at Grahamstown (SZENDREI and MCELHINNY, 1956a). In view of the evidence put forward above for the presence of disturbances due to distortion of the layers, it is considered premature to make any further attempts to derive recombination coefficients and solar models until the effects of the disturbances are better understood. It would appear, however, that these effects may be responsible for the odd shape of the curves of ionization at fixed heights given by SZENDREI and MCELHINNY (1956c, Figs. 3, 6 and 7), and that a revision of their estimates of recombination and attachment coefficients may be possible later.

Acknowledgements—The writer is indebted to the National Telecommunications Research Laboratory of the South African Council for Scientific and Industrial Research, Johannesburg, for arranging to have the ionograms taken at one minute intervals at Cape Town and Johannesburg, and for making the records available for scaling and study. He also wishes to thank Miss J. M. POWRIE, B.Sc., for assistance in the computation of the theoretical curves in Figs. 3–6.

REFERENCES

BONNET G., HUNAERTS J. and NICOLET M. 1957 *Bull. Acad. Roy. Sci. Coloniales* N.S. III, **4**, 964.

MCELHINNY M. W. 1958 Ph.D. Thesis, Rhodes University.

McElhinny M. W.	1959	*J. Atmosph. Terr. Phys.* **14,** 273.
Minnis C. M.	1956	*Solar Eclipses and the Ionosphere* p. 204. Pergamon Press, London.
Minnis C. M.	1958	*J. Atmosph. Terr. Phys.* **12,** 272.
Munro G. H. and Heisler L. H.	1958	*J. Atmosph. Terr. Phys.* **12,** 57.
NTRL	1955	*Bulletin of Ionospheric Characteristics at Johannesburg and Cape Town, December* 1954. National Telecommunications Research Laboratory, Johannesburg.
Szendrei M. E. and McElhinny M. W.	1956a	*J. Atmosph. Terr. Phys.* **8,** 108.
Szendrei M. E. and McElhinny M. W.	1956b	*Solar Eclipses and the Ionosphere* p. 74. Pergamon Press, London.
Szendrei M. E. and McElhinny M. W.	1956c	*J. Atmosph. Terr. Phys.* **9,** 118.
Wadley T. L.	1949	*Proc. Instn. Elect. Engrs.* Pt. III., **96,** 483.

OL.
16
959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 376 to 383. Pergamon Press Ltd. Printed in Northern Ireland

Study of "winds" in the *F*-region of the ionosphere during the unusual days in the IGY-calendar

R. N. Singh* and S. R. Khastgir
University College of Science, Calcutta 9, India

(*Received* 2 *March* 1959)

Abstract—The three-spaced-receiver fading records were taken at Banaras with vertical pulsed transmission at 3·8–4·2 Mc/s during the regular world days, world meteorological intervals following the IGY-calendar and also during the SWI's. The records gave information about the movements of the *F*-region of the ionosphere during these unusual days. The velocity and the direction of the "wind" were obtained by the method of similar fades and the cross-correlation method. The experimental arrangements and the methods of measurements have been described. Usually the velocity of the *F*-region "winds" was found to have high values during the world meteorological intervals. The sudden reversal in the direction of the *F*-region "wind" round the midnight hour was confirmed by the observations taken during the regular world days. The sudden reversal in the direction of the "wind" was found to be followed usually by extremely rapid random fading, indicating turbulence in the ionosphere. The fading records on the unusual days in the IGY-calendar have been classified and fully discussed with reference to the random and steady movements of the irregularities which exist for a short time in the *F*-region.

Introduction

The fading records with the technique of three spaced receivers (Mitra, 1949) were taken at Banaras (30°20′ N, 78°3′ E) during the regular world days and world meteorological intervals following the IGY-calendar and also during the SWI's. These observations were taken during the night hours (from 1800 hours to 0600 hours), as there was considerable electrical noise disturbing the observations during the day-time. The lunar–diurnal variations were also taken from 18 to 23 February 1958. The records with the three spaced receivers were taken with pulsed transmission at 3·8–4·2 Mc/s and gave information about the movements of the *F*-region of the ionosphere. The velocity and the direction of the "wind" were obtained by the method of similar fades or by the method of cross-correlation, as applied to the three-spaced-receiver fading records. The study of the *E*-region "winds" was carried out previously in this laboratory with continuous-wave radio signals by Singh and Prakash (1958).

Experimental Arrangements

(a) *Transmitter.* The transmitter employed was of a self-excited Hartley type with a peak power of about 1 kW having a frequency range from 1·82 to 13 Mc/s. The oscillatory valve used for the purpose was a Philips valve of type 2·5/300. A dipole aerial of suitable length with properly matched transmission line was used with the pulse transmitter. The grid-blocking method was employed for the production of pulses. The pulse duration was of the order of 150 μsec and the pulse repetition frequency was 100–130/sec.

(b) *Three spaced receivers.* Three centre-tapped dipole receiving aerials which were horizontal and parallel to one another were erected at the three corners of a right-angled triangle, the two arms of the right angle, each of the order of 100 m, being in the north–south and east–west directions. Three similar radio receivers,

* Lecturer in Physics, Banaras Hindu University.

VOL. 16 1959

each connected with a cathode-ray oscillograph, were worked with these aerials. Each aerial was kept about 25 ft above the ground to remove unwanted reflections from the same. The output from the centre-tap of each dipole was brought to each receiver by a twin-wire transmission line. Each dipole was a single frequency system. For the reception of different frequencies by each dipole system, the impedance was matched by the special aerial-impedance-matching circuit of each of the three Philips Communication Receivers (Model BX 925 A/00a–01a).

(c) *Recording unit.* The recording unit consisted of three oscillographs (DuMont Type 274A) and a moving-film Cossor Camera, driven by a speed-reduced motor of

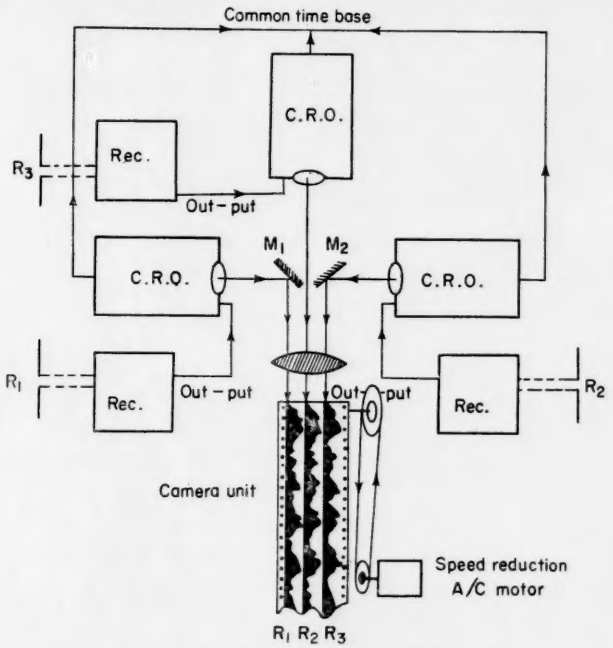


Fig. 1. Block diagram of recording arrangements.

nearly 2 rev/min. The detector output of each receiver was fed into the horizontal deflecting plates of each oscillograph, the vertical plates being earthed. While recording, the oscillographs were kept in such a way that the beam from one of the oscillographs reached the camera directly and the beams from the other two after reflection from two small mirrors, each inclined at an angle of 45° to the direction of either beam. The positions of the mirrors, the oscillographs and the camera were so adjusted that the beams from the spots on the fluorescent screens of the three oscillographs were focused simultaneously on the camera film one above another in succession. The movement of the camera film and the displacement of the focused spots due to signal variations were perpendicular to each other, so that the three continuous fading curves were traced on the camera film, along three parallel horizontal directions.

A block diagram showing the positions of the dipole receiving aerials, receivers, oscillographs, camera unit, etc., is shown in Fig. 1.

Geometry of the triangular arrangement

Let the three horizontal dipole aerials, parallel to one another be situated at the three corners R_1, R_2, R_3 of a right-angled triangle and let the pulse-transmitter be located at T (see Fig. 2). In the case of a single-hop propagation, the pulse from T to be received at R_1 must have been reflected from a point in the ionosphere the projection of which on the ground is R_1' which is half-way between T and R_1. Similarly the pulse from T to be received at R_2 and R_3 must have been reflected from points in the ionosphere the projections of which on the ground are R_2' and

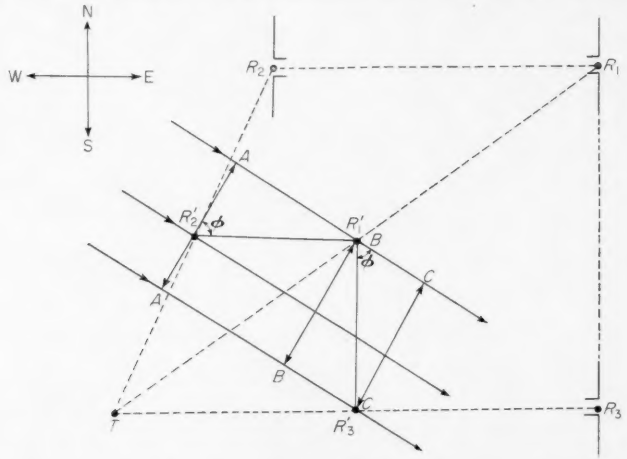


Fig. 2. Geometry of the triangular arrangement in the three-spaced-receiver method.

R_3' which are half-way between the pulse-transmitter at T and the receiving dipoles at R_2 and R_3.

It is evident from the geometry of the triangular system that each of the base-lines $R_1'R_2'$ and $R_1'R_3'$ is half of each of the base-lines R_1R_2 and R_1R_3.

Method of similar fades

The fading patterns where the time-displacements of the maxima of the two records corresponding to the aerials in the north–south or in the east–west directions could be directly measured were analysed by the method of similar fades. The formulae used in this method are developed as follows.

Let an irregularity in the form of an ion-cloud in the ionized layer of an undulation in the lower boundary, having the dimensions of a few wavelengths, travel in a direction making an angle ϕ with the north–south direction. When the irregularity reaches the point of ionospheric reflection, the projection of which on the ground is R_2', the scattered components are increased and the vector sum of these components gives a maximum on the fading record. As the irregularity advances on the direction as shown by the arrow, it crosses the point of ionospheric reflection the projection of which on the ground is R_1', giving a maximum on the other record displaced by the time required by the irregularity to move from AA to BB. In this way, we get maxima on the three fading records corresponding to the positions of the irregularity AA, BB and CC.

VOL. 16 1959

If the velocity of the irregularity in the direction of the arrow be V, then the time-lags between the maxima of the fading records at R_1, R_2 and R_1, R_3 are given by:

$$t_x = \frac{R_1' R_2' \sin\phi}{V} = \frac{R_1 R_2 \sin\phi}{2V} = \frac{x \sin\phi}{2V} \tag{1}$$

$$t_y = \frac{R_1' R_3' \cos\phi}{V} = \frac{R_1 R_3 \cos\phi}{2V} = \frac{y \cos\phi}{2V} \tag{2}$$

where $x = R_1 R_2$ and $y = R_1 R_3$

From these time-shifts obtained from the actual records we can measure the "apparent components", V_x and V_y defined by

$$V_x = \frac{x}{t_x} = \frac{2V}{\sin\phi} \tag{3}$$

$$V_y = \frac{y}{t_y} = \frac{2V}{\cos\phi} \tag{4}$$

Therefore from (3) and (4) we get

$$\frac{1}{4V^2} = \frac{1}{V_x^2} + \frac{1}{V_y^2} \tag{5}$$

and

$$\phi = \tan^{-1}\frac{V_y}{V_x} \tag{6}$$

Thus equations (3), (4) and (5) enable us to calculate the magnitude of the true "wind" velocity from a knowledge of the "apparent" components and equation (6) gives the direction of the "wind" with reference to the north–south direction.

Cross-correlation method

This method of statistical analysis is used to determine accurately the time-displacements between two fading records where they are not clearly discernible. In such cases, the time-shifts are obtained by drawing two cross-correlograms corresponding to the fading records at the receiving sites R_1, R_2 and R_1, R_3 by using the formula for the cross-correlation coefficient:

$$\rho(\tau) = \frac{\Sigma[x(t) \,.\, y(t-\tau)]}{\{\Sigma[x(t)]^2 \,.\, \Sigma[y(t-\tau)]^2\}^{1/2}} \tag{7}$$

where $x(t)$ is the departure from the mean signal amplitude on one record and $y(t)$ on the other at any instant t. The signal amplitude at the interval of τ sec with respect to the time t is noted and the cross-correlation coefficient is calculated with help of formula (7).

The mean signal amplitude $\bar{R}$ is related to the most probable amplitude R_m by the relation:

$$\bar{R} = \left(\frac{\pi}{2}\right)^{1/2} . R_m \tag{8}$$

where R_m is obtained from the actual amplitude distribution curve drawn from the fading records.

The values of the cross-correlation coefficient calculated in this way for different values of τ from one pair of fading records at R_1 and R_2 give one cross-correlogram. Similar calculations from the other pair of fading records at R_1 and R_3 give the other cross-correlogram. The positions of the prominent maxima with respect to the time of reference, where $\tau = 0$, in the two cross-correlograms give the corresponding time-displacements, t_x and t_y required for the determination of the magnitude and direction of the ionospheric "wind".

Experimental Results

The three-spaced-receiver fading records taken on the unusual days in the IGY-calendar can be broadly classified as follows:

(i) Random fading pattern.
(ii) Quasi-periodic pattern showing alternate maxima and minima of an irregular nature.
(iii) Regular periodic pattern persisting for a short time.
(iv) Very rapid random fluctuations.

The most frequently observed patterns were of the random type. A typical set of records showing alternate maxima and minima of an irregular nature is shown in Fig. 3(a). This was taken on 20 February 1958 at 2130 hours IST. The velocity of the F-region "wind" calculated from these records by the method of similar fades was 38 m/sec in direction 17° south of east. Typical regular periodic patterns recorded less frequently are shown in Figs. 3(b) and 3(c). The periodicity in Fig. 3(c) is followed by a steady pattern. These sets of records were taken on 25 March 1958 at 2330 IST and on 21 September 1958 at 0500 IST, respectively. The velocity of the "wind", estimated in the same way was 12 m/sec in direction 35° north of east. Fig. 3(d) represents a set of records showing random fluctuations superimposed on the periodic patterns. This set of records was taken on 23 March 1958 a little after 2400 IST. The estimated velocity of the drift from the relative displacements of the patterns was 30 m/sec in direction 27° north of east.

A typical set of records of extremely rapid fluctuations is shown in Fig. 3(e). The records were taken on 22 September 1958 at 2200 IST. The fading was too rapid to determine the velocity of the ionospheric drift.

The observations during the world meteorological intervals showed very high velocities of the F-region. During the world meteorological interval beginning on 17 March 1958, records were taken each day at 2300 IST for a period of 10 days. Typical records giving high velocity of the "wind", are shown in Figs. 4(a, b and c). From the records shown in Fig. 4(a) and taken on 22 March which was a day of unusual meteoric activity, the velocity of the "wind" as estimated by the cross-correlation method, was 110 m/sec in direction 20° south of east. Fig. 4(b) shows records taken on 19 March at 0530 IST. The velocity of the "wind" estimated in the same way was 76 m/sec in direction 13° south of east. This high velocity was found to persist for a fairly long time. The records shown in Fig. 4(c) and taken

VOL. 16 1959

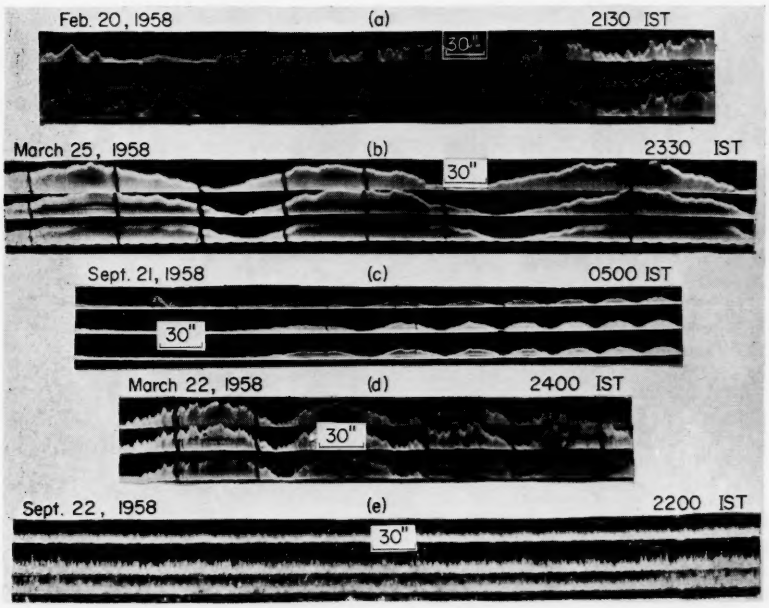


Fig. 3. Three-spaced-receiver fading records taken on the unusual days in the IGY-calendar.
(a) Quasi-periodic pattern showing alternate maxima and minima of an irregular nature (20 Feb. 1958, 2130 IST).
(b) Regular periodic pattern (25 March 1958, 2330 IST).
(c) Regular periodic pattern followed by a steady pattern (21 Sept. 1958, 0500 IST).
(d) Random fluctuations superimposed on regular periodic pattern (22 March 1958, 2400 IST).
(e) Very rapid random fading showing turbulence in the *F*-region.

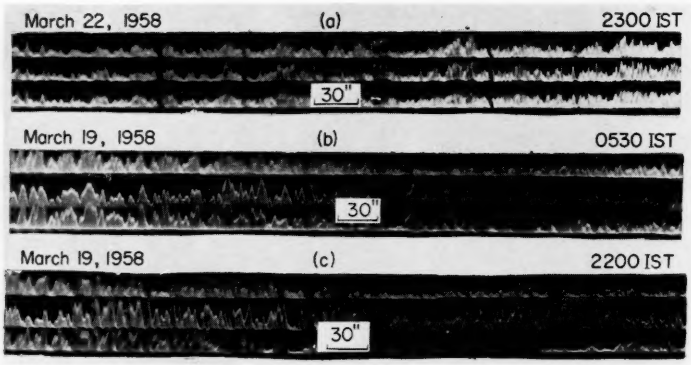


Fig. 4. Random fading records during world meteorological interval.
(a) Recorded on 22 March 1958, 2300 IST.
(b) Recorded on 19 March 1958, 0530 IST.
(c) Recorded on 19 March 1958, 2200 IST.

VOL.
16
1959

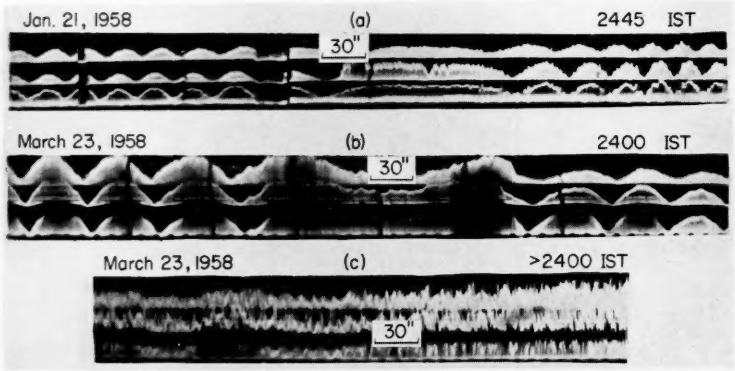


Fig. 5.
(a) Regular periodic pattern showing sudden reversal in the direction of the *F*-region drift (21 Jan. 1958, 2445 IST).
(b) Regular periodic pattern showing sudden reversal in the direction of the *F*-region drift (23 March 1958, 2400 IST).
(c) Very rapid random fading showing turbulence in the *F*-region after the reversal (23 March 1958, a little after 2400 IST).

on the same day after 7½ hr at 2200 IST gave approximately the same velocity and direction of the ionospheric "wind" as at 2200 IST.

Considering all the fading records taken during the regular world days and world meteorological intervals following the IGY-calendar and also during SWI's, the average velocity of the *F*-region drift was found to be 45 m/sec in direction 30° north or south of east before midnight and 20° north or south of west after midnight. In general, the east–west component was found to be in the range 75–180 m/sec which was much larger than the north–south component which was found to be within the range 40–70 m/sec. Sometimes the north–south component was found to have a very low velocity giving a drift mainly in the east–west direction. In such cases, the east–west component was always more than 60 m/sec.

The sudden reversal in the direction of the *F*-region drift round the midnight hour observed by BRIGGS and SPENCER (1954) was confirmed by the observations taken during the regular world days in January and February 1958. It was found that the time for the sudden reversal in the direction of the drift changed only slightly from day to day round the midnight hour. Fig. 5(a) shows the abrupt reversal in the direction of the *F*-region drift observed on 21 January 1958 at 2445 IST. The reversal was found to take place round the long periodicity region in the middle, the time displacements of the maxima being of opposite sign on either side of the region. The values of the velocity, as determined by the method of similar fades, were found to be 21 m/sec in direction 43° south of west before the reversal of the direction and 30 m/sec in direction 37° north of east after the reversal. The records shown in Fig. 5(b) and taken on 23 March 1958 at 2400 IST also revealed a similar reversal in the direction of the *F*-region drift, the estimated values of the velocity being 17 m/sec in direction 19° north of west before the reversal of the direction and 23 m/sec in direction 12° south of east after the reversal. Such reversal in the direction of the drift was usually followed by turbulence in the ionosphere giving rise to very rapid random fading. The records shown in Fig. 5(c) and taken on the same day, a little after the records shown in Fig. 5(b) were taken, revealed such rapid random fluctuations. In many such cases of very rapid fading, it was not possible to apply the cross-correlation method for the evaluation of the velocity of the ionospheric drift. In cases, where the cross-correlation method could be applied, the estimated velocity was low, ranging between 8 and 17 m/sec.

DISCUSSION OF THE EXPERIMENTAL RESULTS

The irregularities, in the form of undulations or unevenness in the lower boundary or in the form of ion-clouds or patches of increased ionization, prevail also in the *F*-layer to some extent, although on account of the higher rate of diffusion due to larger mean free path and higher temperature than in the *E*-layer, the irregularities are not likely to persist for any length of time, as they are diffused away soon after they are formed. The *F*-layer fading patterns, with vertical-incidence pulsed transmission, as recorded simultaneously on the three spaced receivers are attributed to the interference between the wavelets scattered from such transient irregularities and the waves reflected from the "smooth" ionized layer, the resultant amplitude at any instant depending on the phase-difference

between scattered wavelets and the waves reflected from the ionosphere. The nature of the fading pattern caused by such interference must needs be random owing to the random motion of the irregularities. If there is a steady "wind" in the ionosphere and the irregularities move past the receiving site, a periodicity in the fading pattern is expected as the phase-difference between the scattered and the reflected waves would change continuously giving rise to alternate maxima and minima for the resultant amplitude. If there is a large number of irregularities distributed at random, both with regard to space and time, and there is a steady drift of the ionosphere carrying with it the irregularities, we should expect a quasi-periodic pattern with alternate maxima and minima of an irregular nature. Such quasi-periodic patterns were frequently recorded. A typical set of such records has been shown in Fig. 3(a). If however there is a *single* ion-cloud or there are only a *few* irregularities, a steady ionospheric "wind" would give rise to a regular periodic pattern. Such patterns were less frequently observed. We have shown such records of regular periodic patterns in Figs. 3(b) and 3(c) and Figs. 5(a) and 5(b). The periodicity followed by steady amplitude, shown in Fig. 3(c) is similar to the pattern observed by Mitra (1953) who attributes it to the drift of a *single* ion-cloud.

Whenever the random motion of the irregularities was perceptibly large, random fluctuations were found superimposed on the periodic patterns as in the records shown in Fig. 3(c). The sudden reversal in the direction of the ionospheric drift observed during the regular world days was indicative of an approach to turbulence in the ionosphere and was often followed by extremely rapid random fading which was expected in the turbulent state of the ionized region. A typical set of records showing such extremely rapid fading after the sudden reversal in the direction of the F region drift has been shown in Fig. 5(c).

The three-spaced-receiver records taken during the regular world days showed quasi-periodic, regular periodic and random patterns. Calculation by the method of similar fades or by the cross-correlation analysis did not, however, give a high velocity of the F-region drift. Similar fading records taken during the world meteorological intervals usually showed random variations of signal amplitude (see Figs. 4(a, b and c)). During intense meteoric activity, the random motion of the ionospheric irregularities should be large and such large random motion would mask the periodicity caused by the steady drift. The analysis of these random fading records gave distinctly high values for the velocity of the drift which one should expect during the world meteorological intervals.

Conclusion

The investigations on ionospheric drift during the unusual days of the IGY-calendar are still in progress. The results of further studies will be reported in due course.

Acknowledgements—The writers record their thanks to the Council of Scientific and Industrial Research for sponsoring a research scheme on ionospheric movements. They also thank Mr. A. Prakash, M.Sc., for occasional technical help.

VOL. 16 1959

References

Briggs B. H. and Spencer M.	1954	*Rep. Progr. Phys.* **17,** 247.
Mitra S. N.	1949	*Proc. Instn. Elect. Engrs.*, Pt. III, **96,** 441.
Mitra S. N.	1953	*Indian J. Phys.* **27,** 561.
Singh R. N. and Prakash A.	1958	*J. Sci. Res. Banaras Hindu Univ.* **8,** 117.

OL.
16
959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 384 to 392. Pergamon Press Ltd. Printed in Northern Ireland

Turbulence at altitudes of 80–100 km and its effects on long-duration meteor echoes

J. S. Greenhow and E. L. Neufeld
University of Manchester, Jodrell Bank Experimental Station

(*Received* 23 *March* 1959)

Abstract—The two theories of long-duration radio echoes from meteor trails based on scattering from an over-critically dense ionized column, and incoherent scattering from a trail rendered under-dense by small scale turbulence are compared. Examination of the characteristics of enduring meteor echoes shows that the incoherent scattering theory is untenable, if small scale turbulence with a time contact of only 0·4 sec is assumed to be present. All the phenomena observed during the lifetime of a long duration meteor echo are readily explained on the basis of multiple reflections from overdense trails. The time constant and scale of the smallest eddies at heights of 80–100 km are shown to be at least 30 sec and 30 m, respectively.

1. Introduction

Recently Booker and Cohen (1956) have introduced a theory of long-duration meteor echoes based on eddy diffusion in the upper atmosphere. The implications of this theory are of considerable importance in the interpretation of long-duration radio echoes from meteor trails, as it leads to results completely at variance with those deduced from the generally accepted theories of long-duration meteor echoes (Greenhow, 1952a; Kaiser and Closs, 1952). Briefly, the current theories are based on the principle that the electron densities in these trails exceed the critical electron density N_c at the radio wavelength concerned. The ionized column expands owing to molecular diffusion, and when the volume electron density at the centre of the trail falls below N_c, the echo vanishes.

VOL. 16 1959

According to Booker and Cohen eddy diffusion up to 10^4 times more rapid than molecular diffusion quickly reduces the electron density to below the critical density. An echo is then obtained from this under-dense ionized trail from irregularities in electron density. It is the purpose of this paper to present direct experimental evidence capable of deciding which of the two mechanisms is primarily responsible for the behaviour of long-duration meteor echoes.

2. Theories of Meteor Echo Durations and Amplitude Fluctuations

2.1. *Over-critically dense trails*

2.1.1. *Echo duration.* Soon after a meteor trail is formed the radial electron density at radius r assumes a gaussian distribution of the form

$$n_r = \frac{\alpha}{4\pi Dt} \exp(-r^2/4Dt) \tag{1}$$

where α is the linear electron density and t the time after formation of the trail. D is the appropriate diffusion coefficient, in this case the molecular diffusion coefficient determined by the collision cross-sections of meteor ions and atmospheric molecules. For heights of from 90 to 95 km an average value of D is 4×10^4 cm^2/sec (Greenhow and Neufeld, 1955). When $\alpha > 10^{12}$ electrons/cm, reflection

is assumed to take place at the critically dense boundary, and the trail gives rise to a long-duration echo.

The duration of the echo, given by the time when the axial electron density falls below N_e, is

$$T_L = \frac{\alpha}{4\pi D N_e} = \alpha \frac{\lambda^2}{4\pi^2 D}\left(\frac{e^2}{mc^2}\right) \tag{2}$$

As an example, a trail with $\alpha = 5 \times 10^{14}$ electrons/cm would give a 100 sec echo at a frequency of 36 Mc/s, when $D = 4 \times 10^4$ cm^2/sec. The radius of the ionized column after 100 sec given by $r_0^2 = 4Dt$, would be 40 m. From equation (1) the maximum radius of the reflecting cylinder during the lifetime of an echo is $\left(\frac{\alpha}{2{\cdot}7\pi N_e}\right)^{1/2}$ and for $\alpha = 5 \times 10^{14}$ electrons/cm this radius is 20 m.

2.1.2. *The amplitude fluctuations.* When a meteor trail is first formed it behaves as a specular reflector, and the radio echo is obtained from the first Fresnel Zone near the specular point. In the presence of turbulence with a scale of several kilometres the meteor trail will distort, and another part of the trail may become orientated so as to reflect radio waves to the observer. Relative movements of the reflecting centres then produce amplitude fluctuations with a period determined by the r.m.s. turbulent velocity V_1, and radio wavelength. Thus fluctuations will commence when a part of the trail some distance from the initial echoing region has turned through a sufficient angle to present a specular reflecting region (Greenhow, 1950). Although the trail will distort considerably, as long as we are concerned only with meteor echoes which have durations less than the time constant of the smallest eddies, the trail will retain a fairly simple geometric configuration. Some degree of specularity will remain up to this time, after which it will be no longer possible to recognize a simple sinuous pattern. Thus the effect of large scale turbulence would be to give additional discrete echoes at ranges differing from that of the original echo. The times of appearance of these additional echoes will be delayed by amounts depending upon the wind shear and their angular displacement from the initial echoing point.

2.2. *Scattering by electron density irregularities in under-dense trails*

2.2.1. *Echo duration.* In this theory small-scale turbulence with an eddy size $\eta \sim 1{\cdot}3$ m and time constant $t_2 \sim 0{\cdot}4$ sec is assumed to exist in the 80–100 km, region while the large-scale turbulence has an eddy size $L \sim 1{\cdot}6$ km, time constant $t_1 \sim 50$ sec, and r.m.s. velocity $V_1 \sim 35$ m/sec. Up to a time t_2 the molecular diffusion coefficient D is appropriate. After this time eddy diffusion becomes effective, and the eddy diffusion coefficient increases from a value of 4×10^4 cm^2/sec to $V_1 L$ ($\sim 5{\cdot}6 \times 10^8$ cm^2/sec) at $t_1 = 50$ sec. Thus if turbulence of this type is present, the electron densities in even the most dense trails are quickly reduced below N_e and the trail radius increases extremely rapidly (Fig. 1). After only 50 sec, for example, the radius is approximately 2 km.

However, the echo does not disappear when the trail becomes under-dense, because of scattering by inhomogeneities in electron density. As the radius continues to increase the volume electron density decreases further, but the number

of irregularities increases. After the axial electron density falls below N_c, this theory predicts that the echo amplitude should decay in the manner $E \propto t^{-3}$.

2.2.2. *Origin of the amplitude fluctuations.* According to BOOKER and COHEN, small-scale turbulence renders a meteor trail rough in a time $t_2 \sim 0{\cdot}4$ sec. Amplitude fluctuations set in at this time owing to scattering from a large number of small irregularities, and the trail loses its aspect sensitivity. The period of fluctuation would, however, still be determined by the velocity of the largest eddies.

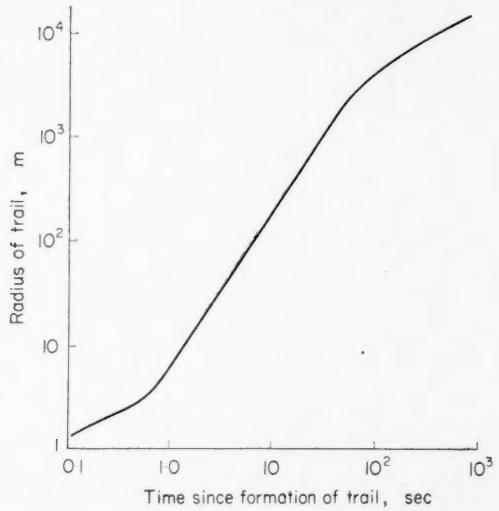


Fig. 1. Radius of a meteor trail as a function of time from formation of the trail, assuming an eddy diffusion coefficient rising from 4×10^4 cm²/sec at 0·4 sec, to $5{\cdot}6 \times 10^8$ cm²/sec at 50 sec (after BOOKER and COHEN).

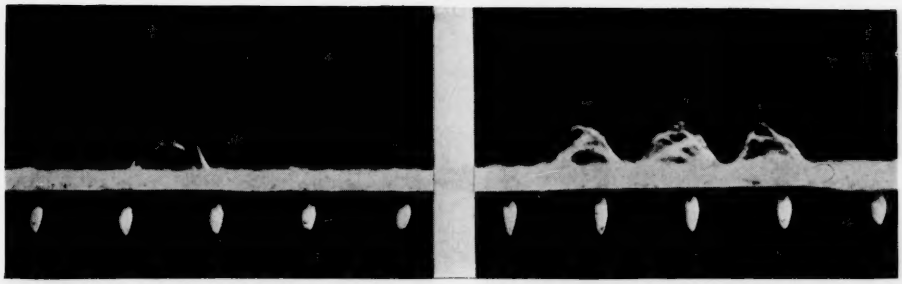
3. EXPERIMENTAL RESULTS

3.1. *Increase in trail radius*

The scattering of radio waves by an inhomogeneous, under-critically dense column of electrons is such that after 50 sec the diameter of the column responsible for scattering the radio waves should be of the order of 4 km (Fig. 1). Thus scattering takes place from irregularities over a depth of 4 km, compared with reflection at a boundary in the over-dense theory. It is therefore only necessary to use a narrow pulse transmitter, and examine the spread in range of the echoes, in order to test the incoherent scattering theory. An experiment of this type has been performed (GREENHOW, 1952b) using a transmitter of pulse length 3 μsec at a frequency of 36 Mc/s.

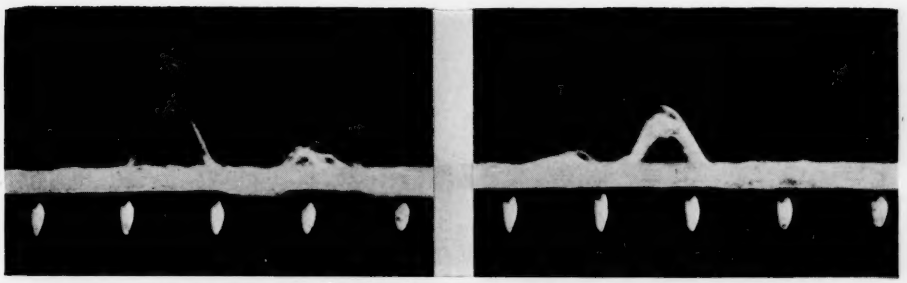
An example of the echoes recorded during the narrow pulse experiments is shown in Fig. 2. An automatic strobe is used to ensure that each echo is displayed on a range time base only 5 km in length, enabling increases in echo width of less than 50 m to be resolved. Successive cine frames of duration 1 sec are taken at intervals of 1·2 sec. At the same time the pulse-to-pulse fluctuations in echo amplitude are recorded on a continuously moving film. Sixty-eight meteor echoes

VOL. 16 1959



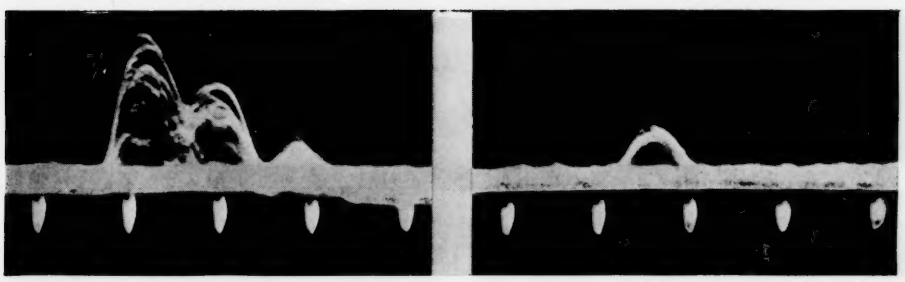
(a) $t = 2{\cdot}4$ sec

(d) $t = 24$ sec



(b) $t = 8{\cdot}4$ sec

(e) $t = 28{\cdot}8$ sec



(c) $t = 15{\cdot}6$ sec

(f) $t = 31{\cdot}2$ sec

Fig. 2. Stages in the lifetime of a long duration meteor echo. The duration of the echo was 32 sec. The photographs are cine frames of a range amplitude display, with 1-km range markers.

VOL.
16
1959

with durations between 10 and 90 sec have been recorded, and measurements of the pulse width made in the last second of their lifetimes (Fig. 2(f)). The increase in final echo width over the width of an echo from a discrete reflector is shown as a function of echo duration in Fig. 3. Echoes of duration 10 sec show a mean terminal excess spread in range of 10 m instead of 300 m required by the turbulence theory. Echoes of 50 sec duration show a small increase in width of only 50 m, compared with the 4 km spread required by this theory. Thus it is quite clear that long-

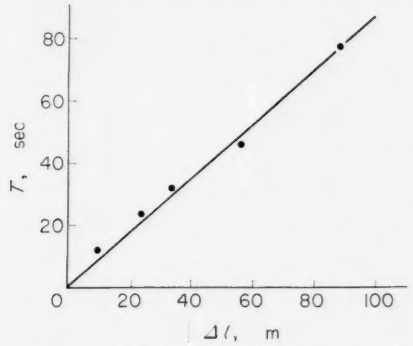


Fig. 3. Increase in echo pulse width (Δl) over the width of an echo from a discrete reflector, measured in the last second of meteor echoes with durations T. Median values of Δl are taken for duration groups.

duration meteor echoes do not show the extremely large increases in pulse width to be expected from a large-diameter partially reflecting trail. Instead, the pulse width at the end of an enduring echo is very little wider than that expected from a perfectly discrete reflector (the significance of the small increase in width actually observed, and the multiple scattering shown in Figs. 2(b) to (e) will be discussed later).

3.2. *Echo structure*

On a large scale turbulence mechanism, the stages in the life of a long-duration echo would be as follows.

(a) A single discrete echo is obtained from an initially linear trail.

(b) The trail begins to distort, and additional discrete echoes appear at other ranges and later times. The overall width of the echo increases, but it shows a simple discrete structure.

(c) Because of the effects of diffusion (and attachment), the electron density along the trail falls, and the discrete echoes begin to disappear as the electron density drops below N_e.

(d) Towards the end of the echo the width of the reflected pulse will once more shrink to that from a discrete reflector, as the length of the over-critically dense section of trail collapses to zero. In the case of a bright meteor giving an echo of many seconds duration, the longest enduring echo is very likely to come from a bright flare along the trail. A flare of only 0·5 magnitude which would hardly be noticed visually, gives a 60 per cent increase in α, and also in duration. Thus the final echoing region may only be of the order of 0·5 km in length.

Stages (a) and (d) are precisely those observed in the lifetime of long-duration echoes when examined with high-range resolution. An example is shown in Fig. 2. After 2·4 sec from the beginning of the echo the pulse is still discrete although a slight broadening has taken place. At times of 6 and 8 sec, additional discrete echoes appear at greater ranges as the trail further distorts, and the overall echo width increases by 2 km. At still later times the discrete reflectors disappear one by one, until after 26 sec only a single narrow echo pulse remains. The most enduring component was the last to appear, although the first was the largest in amplitude. Inspection of many such enduring echoes shows that discrete components may appear with delays of up to 30 sec, showing that the trail still maintains some aspect sensitivity even after this time.

These observations cannot be explained by a theory which requires the trail to be rough after only 0·4 sec, although the discrete echo structure, with delays in appearance of many seconds, is readily explained on a large-scale turbulence theory.

3.3. *Time taken for the onset of fluctuations*

Measurements of pulse shape show that when fluctuations commence within the first 0·5 sec of a specular echo, the increase in echo width due to the appearance of an additional reflector is approximately 10 m. This enables the separation of the reflecting centres along the trail to be determined, and gives an estimate of the angular distortion θ; thus for a time of 0·5 sec, $\theta \sim 0{\cdot}01$ rad. A wind shear of 20 m-sec^{-1}-km^{-1} is required to turn a length of trail through this angle in 0·5 sec and direct measurements show that there is a 25 per cent probability of finding a shear of this magnitude over any 0·5 km height difference in the meteor region (GREENHOW and NEUFELD, 1959). This probability must be approximately doubled as distortion may occur on either side of the initial reflector. There is, therefore, almost a 50 per cent probability that a meteor echo will begin to fluctuate in less than 0·5 sec through the effect of large-scale turbulence. This is in good agreement with the experimental observations (GREENHOW, 1952a).

The median value of turbulent wind shear in the meteor region is 10 m-sec^{-1}-km^{-1}, and for this value a section of trail 18° from specular would take up to 30 sec to distort sufficiently to give an echo. As bright meteor trails can be longer than this, large-scale distortion adequately accounts for the observed delays of up to 30 sec in the appearance of additional discrete echoes. Similarly the radio echo from a visual meteor trail which does not initially present a specular reflecting region, may be delayed in appearance by many seconds (MCKINLEY and MILLMAN, 1949).

As the condition of specularity is reached at the same time irrespective of wavelength, no significant differences would be expected in the times of onset of fluctuations at different wavelengths. Slight differences would arise because of uncertainty in determining the time at which fluctuations commence, because of differences in the re-radiated polar diagrams of the scattering centres. This is in agreement with previous observations (GREENHOW, 1952a). The above results therefore show that the time constant of the small eddies must be at least greater than 30 sec as echoes cannot be delayed in appearance by a time greater than t_2.

VOL. 16 1959

3.4. *The amplitude decays of long duration meteor echoes*

3.4.1. *The echo duration.* Evidence has been presented showing that the behaviour of long-duration meteor echoes can be explained only by an over-dense trail mechanism, with large-scale turbulence. The amplitude decays of long duration echoes will now be examined, as it was upon a supposed t^{-3} decay that the verification of the small-scale turbulence theory was based.

Assuming this theory, any meteor trail of reasonable electron line density ($\alpha < 10^{15}$/cm) must fall below N_e in about 5 secs at $f \sim 30$ Mc/s. In order to remain over-dense for 100 sec, the impossibly high linear electron density of 10^{19}/cm is required (higher still with attachment).

Booker and Cohen analysed about fifty long-duration echoes and found that 5 sec was about the time at which the majority began to decay. However, by observing for a sufficient length of time, echoes which do not begin to decay for over 100 sec may be observed. Correlation of visual meteor brightness with radar echo duration shows that a 100 sec echo is produced by a -5 magnitude meteor (Millman, 1950). Because of the meteor mass distribution it is therefore necessary to observe for many days before detecting a 100-sec echo at 30 Mc/s (except during intense shower activity). Such echoes are not due to abnormal conditions, and observations of the amplitude fluctuations indicate that the r.m.s. turbulent wind speeds are of the usual magnitude. Also the echo amplitudes are only of the same order as those of a few seconds duration, showing that the echoes are not being seen further down the amplitude decay curves. Thus the existence of such echoes is not compatible with the presence of rapid eddy diffusion in the meteor region.

3.4.2. *Amplitude decays.* The amplitude decays of fifty echoes with durations greater than 20 sec at $f = 36$ Mc/s, some of which have corresponding echoes at 72 Mc/s, have been analysed. Ten of the echoes had durations between 100 and 300 sec, and log amplitude–log duration curves for a number of these are given in Fig. 4. Inspection of such diagrams shows that the echo amplitudes remained substantially constant over much of their lifetimes, before decaying into the receiver noise. For example, on the average the ten echoes lasting over 100 sec were greater than 0·5 their initial amplitudes for 75 per cent of their lifetimes, and greater than 0·25 of this amplitude for 88 per cent of the time.

In general no fixed power-law decay could be fitted to the final decay phases (Fig. 4). This also applies to the decays of some echoes in the shorter-duration range from 1 to 20 sec, which were also examined. As a rule the final disappearance was fairly sudden, although these long echoes could persist at a small steady mean amplitude for many seconds. Often the decay is quite irregular, as would be expected for an echo consisting of a number of discrete reflectors which disappear at different times.

Initial echo amplitude is calculated assuming a linear trail, and the echo is effectively obtained from half the first Fresnel zone (approximately 750 m in length at $f = 36$ Mc/s.) When the trail distorts, this simple picture no longer holds as additional reflecting centres appear. These additional reflectors are comparatively small in number (less than six), and assuming random phases, would only double the mean echo amplitude. Towards the end of the echo when the length

12A—(4 pp.)

of the over-dense trail shrinks to the order of 1 km, the echo amplitude will decay, as we now have to compare the signal from an irregular reflector with that from a smooth reflector of the same length which scatters in phase. It is not profitable to attempt to calculate the precise echo amplitude for such a meteor trail, as extreme

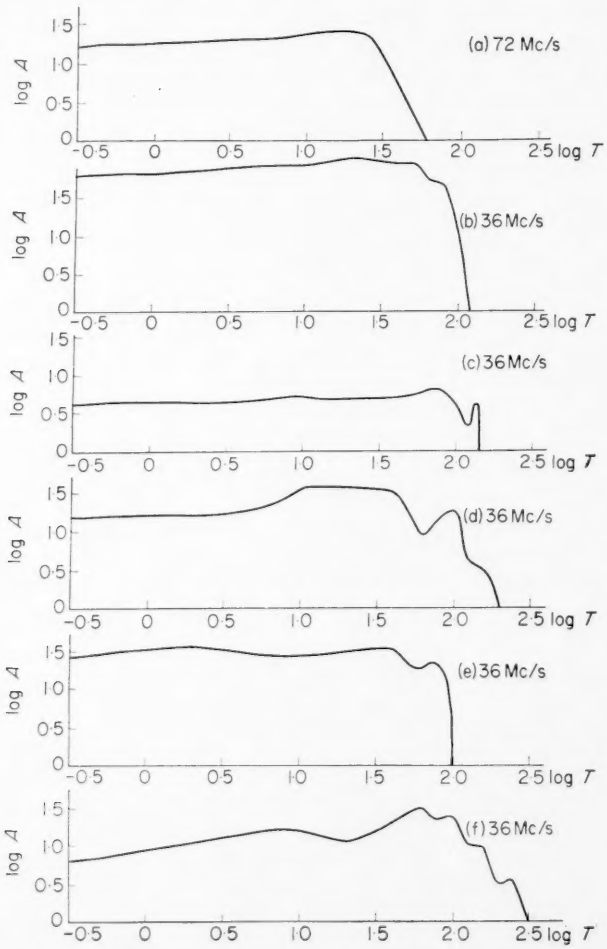


Fig. 4. Log amplitude (A)–log time (T) decay curves of long duration meteor echoes; (a) and (b) are an echo pair at 72 and 36 Mc/s.

simplifying assumptions would have to be made. It has been shown that the echo amplitude will not differ greatly from that of a smooth trail, and this view is supported by experimental observations discussed here, and by the observations of MCKINLEY (1953).

The reasons for the failure of small scale turbulence to affect significantly meteor echoes of many seconds duration will now be considered.

VOL. 16 1959

4. Discussion

4.1. *Eddy size and the turbulence power*

Theories of homogeneous turbulence show that in any turbulent system there should exist a whole range of eddy sizes from the largest, of scale L, r.m.s. velocity V_1 and decay time t_1, to the smallest of scale, η, velocity V_2 and decay time t_2. Energy is supplied to the turbulence at a rate ε erg-g^{-1}-sec^{-1}. This energy is finally dissipated in the form of heat, when viscous forces, which limit the size of the smallest eddies, become significant. Equations of the form $L = V_1 t_1$ and $\eta = V_2 t_2$ relate the scales, velocities and decay times of eddies of any size. The turbulent power is given by

$$\varepsilon = \frac{V_1^2}{t} = \frac{V_1^3}{L} \tag{3}$$

and for the scale and decay time of the smallest eddies (Batchelor, 1953) we have

$$\eta = \left(\frac{\nu^3}{\varepsilon}\right)^{1/4} \tag{4}$$

$$t_2 = \left(\frac{\nu}{\varepsilon}\right)^{1/2} \tag{5}$$

or

$$t_2 = \frac{\eta^2}{\nu} \tag{6}$$

The value of t_2 is of considerable importance, as it determines the time when eddy diffusion becomes effective; ν, the kinematic viscosity, is easily determined from kinetic theory considerations and is equal to the molecular diffusion coefficient (this is greater than the diffusion coefficient for meteor ions, as positive ions have a greater collision cross-section than neutral molecules).

Booker and Cohen derive a value for the time constant of the largest eddies by a manipulation of the Richardson number Ri. They assume that the condition that turbulence should not decay in the 80–100 km region is $Ri < 1$, and arrive at a value $dV/dz = 20$ m-sec^{-1}-km^{-1} for the wind gradient. They then interpret $1/dV/dz$ to be identical with t_1. Taking $V_1 = 35$ m/sec, this gives the extremely high value of the turbulence power $\varepsilon = 2{\cdot}5 \times 10^5$ erg-g^{-1}-sec^{-1}—some 10^5 times higher than the troposphere value. Equations (5) and (4) then give 0·4 sec and 1·3 m as the time constant and scale of the small eddies.

However, it is in the use of Richardson's criterion that the difficulty arises, and this problem will be discussed in the light of recent measurements of turbulence in the 80–100 km region (Greenhow and Neufeld, 1959). Although turbulent wind shears as high as 20 m-sec^{-1}-km^{-1} with r.m.s. velocities of 25m/sec are found, the shears associated with the average winds are less than 2 m-sec^{-1}-km^{-1}.

Even on occasions when the uniform wind gradients are zero, r.m.s. turbulent velocities of 25 m/sec are still present. Furthermore, even when the steady drifts themselves are zero, the turbulent velocities are not reduced. Thus turbulence in the meteor region is not associated with existing wind speeds or wind gradients, and Richardson's criterion cannot be applied.

4.2. *Nature of the turbulence at heights of from* 80–100 *km*

The large-scale turbulence at heights of from 80 to 100 km is found to be distinctly anisotropic. Although the vertical scale of the largest eddies is approximately 7 km, the horizontal scale is of the order of 150 km. The measured time scale is approximately 6×10^3 sec and r.m.s. velocity 25 m/sec. A relationship of the form $L = V_1 t_1$ thus holds for $t \sim 6 \times 10^3$ sec, $V_1 \sim 25$ m/sec and $L \sim 150$ km.

Equation (3) now leads to a value of $\varepsilon \sim 10^3$ erg-g^{-1}-sec^{-1} and for $\nu = 2 \times 10^5$ cm^2/sec, (5) gives for the scale of the smallest eddies

$$\eta = 15 \text{ m}$$

This is in reasonable agreement with the value of 50 m deduced from correlation curves (GREENHOW and NEUFELD, 1959) and the geometric mean $\eta \sim 25$ m has been taken as the best available figure for the scale of the smallest eddies.

Adopting this value of η, equation (6) now gives $t_2 \sim 30$ sec as the time constant of the smallest eddies (compared with 0·4 sec used by BOOKER and COHEN). Evidence that t_2 is at least greater than 30 sec is given by measurements of delays in the appearance of discrete echoes (Section 3.3). Similarly, the small increases in width of approximately 30 m at the ends of echoes with durations of about 30 sec (Fig. 3) set an upper limit to the scale of the smallest eddies. This spread may be explained if the short length of critically dense trail has deformations with a depth of 30 m. Thus for $t_2 \sim 30$ sec, η must be of the order of 30 m. However, because of the unusual nature of the large eddies, it is not certain whether theories of homogeneous turbulence can be applied to the 100 km level, and it is possible that even these small eddies do not exist.

In view of the large minimum value now derived for the time scale of the smallest eddies, it is clear why small-scale turbulence does not significantly affect the characteristics of echoes with durations as high as 100 sec. Thus all the phenomena observed during the lifetime of long-duration meteor echoes can be simply explained in terms of scattering from over-critically dense columns of ionization, modified by the effects of large-scale turbulence.

VOL. 16 1959

REFERENCES

BATCHELOR G. K.	1953	*Homogeneous Turbulence.* Cambridge University Press.
BOOKER H. G. and COHEN R.	1956	*J. Geophys. Res.* **61,** 707.
GREENHOW J. S.	1950	*Phil. Mag.* **41,** 682.
GREENHOW J. S.	1952a	*Proc. Phys. Soc. Lond.* (B) **65,** 169.
GREENHOW J. S.	1952b	*J. Atmosph. Terr. Phys.* **2,** 282.
GREENHOW J. S. and NEUFELD E. L.	1955	*J. Atmosph. Terr. Phys.* **6,** 133.
GREENHOW J. S. and NEUFELD E. L.	1959	*J. Atmosph. Terr. Phys.* In press.
KAISER T. R. and CLOSS R. L.	1952	*Phil. Mag.* **43,** 1.
MCKINLEY D. W. R.	1953	*Canad. J. Phys.* **31,** 758.
MCKINLEY D. W. R. and MILLMAN P. M.	1949	*Proc. Inst. Radio Engrs., N.Y.* **37,** 364.
MILLMAN P. M.	1950	*J. R. Astr. Soc. Can.* **44,** 209.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 393 to 395. Pergamon Press Ltd. Printed in Northern Ireland

RESEARCH NOTES

Escape of helium

(*Received* 16 *June* 1959)

THERMAL diffusion influences the altitude distribution of helium atoms and hence the flux escaping from the exosphere. In a recent paper (BATES and MCDOWELL, 1957) an attempt was made to allow for the effect. The formulae which were presented are correct, but

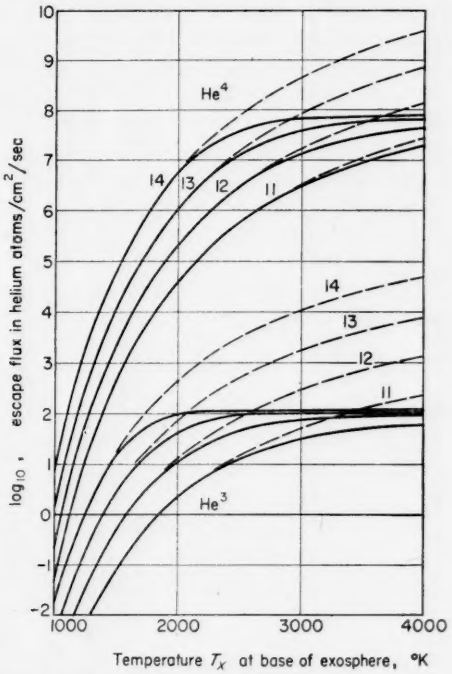


Fig. 1. The broken curves give log l and the full curves give log l' where l and l' are helium escape fluxes, the former obtained ignoring the finiteness of the diffusion coefficient and the latter obtained taking account of the finiteness of this coefficient. The upper set refers to He^4 and the lower set to He^3. The value of log N_D is shown on each curve, N_D being the number density of air atoms and molecules per cubic centimetre at the level at which diffusive separation is assumed to begin. On the model of the atmosphere adopted, the following are corresponding values of N_D and T_D where T_D is the temperature at this level: 10^{11}/cm^3, 370°K; 10^{12}/cm^3, 280°K; 10^{13}/cm^3, 220°K; 10^{14}/cm^3, 190°K.

unfortunately the thermal diffusion factor involved was taken to be positive in the computations, whereas of course it is in fact negative. We are indebted to Dr. F. S. JOHNSON for drawing our attention to the error which we now correct.

Fig. 1 gives the re-computed escape fluxes. They are considerably greater than those given originally.

In examining the results it should be borne in mind that the rates at which He^4 and He^3 atoms are released into the atmosphere are estimated to be about 2×10^6/cm^2-sec

and 2/cm²-sec, respectively. Reservations regarding these estimates are made in the earlier paper.

Rocket measurements indicate that diffusive separation sets in at a lower level than expected by most theorists.* However the number density N_D of air atoms and molecules at the level concerned can scarcely be much greater than 10^{13}/cm³ as supposed by BATES and MCDOWELL. In this case it may be seen that for the escape flux $l'(He^4)$ to be 2×10^6/cm²-sec the temperature T_X at the base of the exosphere would have to be about 2100°K. This is rather higher than studies of the thermo-balance in the upper atmosphere suggest is likely. Moreover, such a value of T_X would lead to the escape flux $l'(He^3)$ being about 60/cm²-sec which is very much greater than can readily be accepted in view of the estimated release rate of 2/cm²-sec. The conclusion (which was also reached by NICOLET, 1957), that there is no *steady* value of T_X at which both isotopes of helium escape at the rates apparently required, is unaltered.

Consider next the hypotheses of SPITZER (1949) that the He^4 escape occurs during brief intervals when the solar ultra-violet emission is exceptionally high. To favour the hypothesis we suppose as before that (i) the heating is world-wide instead of being confined to the sunlit hemisphere; (ii) the brief intervals concerned occur throughout the sunspot cycle and cover 5 per cent of the time rather than 2 per cent as SPITZER contemplated; (iii) the value of N_D is 10^{13}/cm³ even though sudden temperature changes might be expected to promote mixing; (iv) the outward flow is too transient to affect the altitude distribution of the He^4 atoms (but is sufficiently prolonged to affect that of the He^3 atoms there by keeping their rate of loss from exceeding the permissible limit). On these suppositions T_X must rise to about 2800°K for the *average* He^4 escape flux to be 2×10^6/cm²-sec. This is much lower than the corresponding peak temperatures originally deduced but is nevertheless higher than can readily be accepted (cf. BATES and MCDOWELL, 1957). Thus the main argument against SPITZER's hypothesis, though weakened, is still quite strong.

The other matters discussed in the earlier paper do not require comment.

The Queen's University of Belfast — D. R. BATES
and
Royal Holloway College, Englefield Green, Surrey — M. R. C. MCDOWELL

VOL. 16 1959

REFERENCES

BATES D. R. 1959 *Proc. Roy. Soc.* A. In Press.
BATES D. R. and MCDOWELL M. R. C. 1957 *J. Atmosph. Terr. Phys.* **11,** 200.
NICOLET M. 1957 *Ann. Géophys.* **13,** 1.
SPITZER L. 1949 *The Atmospheres of the Earth and Planets* (Edited by KUIPER G. P.), p. 213. University of Chicago Press.

* A brief survey of the evidence has been given by BATES (1959).

E-region winds

(*Received* 22 *June* 1959)

THE number of published observations of *E*-region winds is not large, and the origin of the winds is not completely understood. If one assumes that GREENHOW and NEUFELD's results are representative, it is possible to develop one or two suggestions that may be useful.

The resonance theory of atmospheric solar tides may now be taken as well established (WILKES, 1949, p. 65) and it is usually extended to horizontal movements. The exciting force for the observed oscillations is not well established, but GREENHOW and NEUFELD (1955, p. 560) and WILKES (p. 69) consider the results that would follow if the sun heated the *E*-region. Let us take the hypothesis a little further, and advance some ideas that are supported by the experimental results of the GREENHOW and NEUFELD reference.

If the *E*-region is heated by the sun, there will be a net tendency for air to leave the sub-solar point, which must be treated as an area of indeterminate size. (If, on the other hand, ionospheric winds are due to the sun's gravitational field, there will be a net tendency for air to approach the sub-solar point. GREENHOW and NEUFELD's results favour the heating hypothesis.) The time at any particular place when the wind flows due north, will be later than true noon because it will take some time for the wind, starting from the sub-solar point, to reach the observer's position. The NS semi-diurnal delays show a greater delay in winter, when the sun is further south, than in summer; September and October are anomalous. As the sun passes to the west, the EW component of the local wind will flow first to the east, then to the west.

During the winter the time during which the sub-solar area is active will be short. If it is short compared with 12 hr, the semi-diurnal oscillation will be readily excited. During the summer the sub-solar area will be active for a longer time and the semi-diurnal oscillation will be less, and the diurnal more, readily excited.

The winds so far discussed are elastic—they represent a mass of air that moves a certain distance and comes back again. The runs of wind, computed from GREENHOW and NEUFELD's results are: 178 km to the north once a day and back, and 200 km to the north twice a day and back; to the east 71 km once a day and back, and 123 km twice a day and back.

So far the *E*-region has been considered as a stratum, in which air circulates without loss, as a result of the absorption of the sun's energy within the region itself. The prevailing wind in the NS direction alters this picture. Its value is not overwhelming, for the mean winds from the annual figures are: NS prevailing, 7·1 m/sec, diurnal 4·1 m/sec, semi-diurnal 9·2 m/sec; EW prevailing 4·8 m/sec, diurnal 1·7 m/sec, semi-diurnal 5·7 m/sec. One interpretation that can be advanced is that the air under the sun rises from the *E*-region and escapes, and that roughly once a fortnight the whole *E*-region has come in near the auroral zones and has gone out in the tropics.

During the part of the day when the wind has a southerly component it has, for most of the year, also an easterly one (GREENHOW and NEUFELD, Fig. 3). Because of the escape upwards in the tropics the integrated southerly is greater than the northerly. This will lead to a prevailing wind to the east. Under the path of the sun there cannot be a NS component; instead the air here is rising. This is consistent with the results reported by SKINNER, HOPE and WRIGHT (1958) from Nigeria.

The ideas put forward in this note are obviously too elementary to be accepted without considerable examination, for they are based on very little data, and depend on such gross approximations as the identification of air as a whole with the ionized part of it. It is, however, hoped that their examination may be productive.

Dominion Physical Laboratory
Lower Hutt, New Zealand

W. H. WARD

REFERENCES

GREENHOW J. S. and NEUFELD E. L. 1955 *Phil. Mag.* **46,** 549.

SKINNER N. J., HOPE J. and WRIGHT R. W. 1958 *Nature, Lond.* **186,** 1363.

WILKES M. V. 1949 *Oscillations of the Earth's Atmosphere.* Cambridge University Press.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 16, pp. 396 to 400. Pergamon Press Ltd. Printed in Northern Ireland

Book reviews

D. PEDOE: **Circles:** Pergamon Press Ltd., London, 1958. 78 pp., 20s.

THIS slim volume, which is very clearly written, contains a useful collection of properties of the circle not otherwise obtainable except by consulting a variety of textbooks on a variety of subjects, and can be strongly recommended for any class library.

Chapter I deals with familiar matters such as the nine-point circle, Feuerbach's theorem, inversion and coaxial circles.

The contents of Chapter II will be new to most readers. Since the standard equation of a circle, namely, $x^2 + y^2 + 2gx + 2fy + c = 0$, contains three parameters, a circle can be represented uniquely by a point in three-dimensional space, and this representation is sometimes very illuminating. For example, a family of coaxial circles is represented by a straight line, and, since this line is uniquely determined by two points on it, a coaxial family is determined by any two of its members. After giving several applications of this representation, the chapter concludes with a section on anallagmatic curves, i.e. curves which are self-inverse in a given circle.

Chapter III begins with a résumé of the geometrical representation of complex numbers and then passes on to a discussion of Möbius (or bilinear) transformations which have the characteristic property of transforming circles (in the complex plane) into circles. These transformations are then used to study Poincaré's model of a hyperbolic non-euclidean geometry, the existence of which shows that the parallel postulate of euclidean geometry is independent of the other postulates and cannot be deduced from them.

Chapter IV deals with the problem of showing that of all plane closed curves of given perimeter the circle encloses the greatest area, and is probably more difficult than the rest of the book in the sense that it contains some relatively advanced real variable theory. However, the chapter, like the earlier ones, is very readable and will repay careful study.

The book is obviously not planned to cover any particular examination syllabus, but it nevertheless provides excellent supplementary reading for all students of mathematics from the sixth form onwards. The only thing wrong with the book is its price; £1 seems expensive for a book of 78 pages bound in stiff paper covers.

D. MARTIN

Dictionary of Physics and Allied Sciences Vol. I, German/English. Edited by CHARLES J. HYMAN. Frederich Ungar, New York, 1958. 671 pp., $9.00.

A REVIEWER of a text-book can, and should, read every word of it before he is fully qualified to speak of its merits and demerits. But a reviewer of a dictionary can scarcely be expected to do the same; he can only speak from his experience as a sampler. That is what I have done in this case.

This German–English Dictionary of Physics and Allied Sciences relates to a field of science in which recent progress has been phenomenally rapid. For the English reader it will be of inestimable use for reading and translating German scientific texts, one valuable feature being that it includes sample phrases as well as words. Fortunately, as the title indicates, the Dictionary covers not only the subject of Physics itself, but also the allied subjects of mathematics,

VOL. 16 1959

astronomy, mineralogy, geology, geophysics and meteorology; while, in the field of technology, terms used in electronics, radio, television, telephony, telegraphy, navigation, aviation and ballistics are also included.

One can only hope that the success of the present volume will be such as to encourage the publishers to organize the preparation of a companion Russian–English dictionary!

5th November, 1958 E. V. APPLETON

W. SULLIVAN: **The International Geophysical Year.** No. 521, of International Conciliation, 1959. Carnegie Endowment for International Peace, 1959. 78 pp. 25c.

THE author of this particular issue of International Conciliation, W. SULLIVAN, is a science journalist, on the staff of the *New York Times*, whose work is greatly esteemed in professional scientific circles for its lucidity and accuracy. There should therefore be no one better fitted than he to tell the story of the IGY to the layman. Much of that story, so far as the scientific programme of the IGY is concerned, will of course be familiar enough to readers of this Journal. They may not, however, be so familiar with the manifestations of international rivalry, displayed in the course of the planning and operational phases of the IGY, which are given so much prominence here. One whole chapter is, for example, devoted to the dispute between the two Chinas as regards IGY participation, as a result of which the Chinese People's Republic was the one nation among sixty-seven to withdraw. The chapter on Antarctica likewise deals much with clashes of national claims and interests; while the chapter on earth satellites instances various U.S.S.R. failures to implement data-exchange programmes in this field, and so on.

To the many thousands of geophysicists who have been quietly absorbed in the implementation of the IGY programme, this emphasis on discord rather than concord, by a distinguished commentator, will be rudely awakening. Yet, even here, the great majority of the scientists are seen to have risen above jealousy and emulation.

E. V. APPLETON

M. GOODRICH: **Gravity and Interplanetary Travel.** Vantage Press, New York, 1958, 72 pp. $2.50.

THIS is not a book to be recommended. In his Foreword the author states that, after reading it. anyone "may judge for himself whether interplanetary travel and space ships can be successful". But, unfortunately, the title of the volume is misleading and its contents bear only remotely on the solution of the problems of space travel. The text is, in fact, an amateurish sketch of atomic physics. In his opening paragraph the author states that "There are three important stages of gravity: the pulling power, the repelling power and the overlapping power". This is a fair indication of the standard of the rest of the volume.

E. V. APPLETON

Handbuch der Physik. Geophysik II Vol. 48. Edited by J. BARTELS. Springer, 1957. 1046 pp. £16.18s.6d.

THE second of the three volumes of the *Handbuch der Physik* discusses the physics of the atmosphere, including meteorology, and also contains articles on earth tides and physical volcanology. The articles are not only authoritative but are readable and therefore it is a relevant criticism that the book is too heavy to make continuous reading the pleasure it would otherwise be.

ELIASSEN and KLEINSCHMIDT have written a very full article on dynamic meteorology and anyone wishing to become informed about the problem of the general circulation could hardly have a better guide than this article. The interesting problem of the tides is dealt with in two articles, the first on the tidal forces by BARTELS and the second on the tides of the solid earth by TOMASCHEK. The recent interesting work in Japan on using changes in the earth's magnetic field to give information of developments before and during volcanic activity is described by SAKUMA and NAGATA in their article on "physical volcanology". Vision through the atmosphere and polarization of skylight are described by MIDDLETON and SEKERA, respectively. SVERDRUP gives a general account of the physics of the oceans while other articles deal with waves and swells. An article by LOVELL on meteors is very informative.

The above gives some idea of the coverage of the volume and the concise and intellectually satisfying nature of the writing ensures that the work will be of great importance to the development of geophysics.

S. K. RUNCORN

J. A. RATCLIFFE: **The Magneto-Ionic Theory and Its Application to the Ionosphere.** Cambridge University Press, 1959. x + 206 pp., 40s.

DISCUSSION of the propagation of radio waves in the ionosphere is complicated by the presence of the Earth's magnetic field. An imposed magnetic field renders an ionized gas doubly refracting, so that propagation is different for different wave frequencies and for different directions relative to the magnetic field. Each of the doubly refracted waves has an elliptical polarization that also depends upon the wave frequency and direction of propagation. Formulae to describe these phenomena first made their appearance towards the end of the 1920's. During the ensuing decade they were analysed in some detail and applied to the ionosphere, particularly to the problem of vertical sounding of the ionosphere by radar methods. With the awakened interest in the ionosphere created by the availability of space vehicles, a need exists for a book that integrates the scientific literature on the magneto-ionic theory of radio wave propagation. This book has now made its appearance, written by an outstanding educator and research worker who actively participated in the development of the theory. In writing the book he has not only presented with characteristic clarity the principal features of the work as it was done during the 1930's, but in addition has presented more recent developments of his own not previously published. In particular, he describes in some detail how the dispersion and absorption of a radio wave in a magneto-ionic medium may be understood in detail by examining the waves scattered by individual electrons and combining them with the incident wave to produce what is usually described as the refracted wave.

This book will be of particular interest in universities where a graduate course is taught on subjects connected with the outer atmosphere of the Earth. A minor difficulty that occurs in such courses is that many of the students have been brought up to use rationalized m.k.s. units whereas most of the original literature was written in unrationalized electrostatic and electromagnetic c.g.s. units. Mr. Ratcliffe has largely avoided this difficulty by presenting his equations in such a form that they are true in any consistent system of units whatever. There will, however, be some regret that he did not take the further step of rationalization.

H. G. BOOKER

DAN PEDOE: **The Gentle Art of Mathematics.** Macmillan Company, New York, 1959. 143 p. $3.50.

NO ONE could pretend that this little volume has anything special to say to the atmospheric and terrestrial physicist. But enjoyment should be shared where possible; and, to that end, Professor PEDOE's volume is warmly recommended. The Professor of Mathematics at Khartoum

VOL. 16 1959

clearly professes his subject because he is enthusiastic about it himself. Wearing his own learning lightly, he describes in lively fashion some of the highlights of modern mathematics, including: chance and choice and the law of averages; extra-sensory perception; logic and automatic thinking; topology and geometry; symmetry and series; rational and irrational numbers; and the problem of infinity. As might be expected, the longest chapter deals with mathematical games. It is all highly diverting; and, of course, richly informative as well.

E. V. APPLETON

Contributions in Geophysics in Honor of Beno Gutenberg. Edited by H. BENIOFF, M. EWING, B. F. HOWELL JR. and F. PRESS. Pergamon Press, London, 1958. 244 pp., 60s.

BENO GUTENBERG has taken a leading part in the development of seismological theory, and has made great contributions to present knowledge of the physics of the earth. This volume (Vol. 1. of the International Series of Monographs on Earth Sciences) consists of seventeen papers assembled to pay honour to him on his retirement from the Directorship of the Seismological Laboratory of the California Institute of Technology in 1957.

The papers inevitably differ greatly in depth and importance, and are not connected, but they give up to date information over a wide field of geophysics, as the following summary shows.

M. BÅTH computes the energy of body waves and surface waves for some shallow focus earthquakes, and obtains empirical relationships between the total energy and the energy of Rayleigh waves. P. BYERLY and J. DENOYER compute the work done at the time of faulting by stresses due to elastic deformation in three American earthquakes. R. STONELEY examines the variation of amplitude and energy with depth in Love waves, computing numerically for a model three-layer halfspace.

P. CALOI discusses the relation between earth tilting and subsequent shocks in the Tolmezzo records. W. HILLER writes on the general question of the mechanics of earthquakes.

J. H. HODGSON considers displacements in North-West Pacific earthquakes in the light of Russian fault-plane solutions which make use of transverse wave pulses. C. TSUBOI presents a large amount of statistical data on Japanese earthquakes.

K. E. BULLEN summarizes evidence on the solidity of the earth's inner core. I. LEHMANN discusses the wide scatter of PP, SKS and PS arrival times at epicentral distances 105° to 115°.

J. P. ROTHE reports observations made in Western Europe of refraction and reflection arrivals from explosions in the Alps in 1956 and compares results with those of earlier explosions. H. E. TATEL and M. A. TUVE discuss observations at 1 km depth of local noise and also of pulses from a deep shot. F. BIRCH summarizes experimental data on velocities of elastic waves at high temperatures and pressures, and compares the results with seismic data on velocities in the continental crust.

C. L. PEKERIS and H. JAROSCH present a systematic discussion of the free oscillations of the earth, using conventional analytic, variational and computational methods. F. A. VENING MEINESZ discusses the geophysical history of a geosyncline, assuming pseudo-viscous convection currents at depth. W. A. HEISKANEN and U. A. UOTILA derive new values for the parameters in the formula for variation of gravity with latitude. They find no systematic variation with longitude. H. E. LANDSBERG summarizes methods now in use of data processing in geophysics. W. ELSASSER and W. MUNK show that the displacement of the earth's pole of rotation due to motion of the core is negligible.

Most of the authors assume that the reader is familiar with the subject under discussion. The volume takes too much for granted to form a suitable introduction to geophysics, while it

is not complete enough to be a useful compendium of present knowledge of the subject. It can be recommended as a stimulating picture of a remarkable range of problems, of value in showing some of the points on which experts in the subject are at present concentrating their thought.

E. R. LAPWOOD

T. G. MEHLIN: **Astronomy**. John Wiley, New York; Chapman & Hall, London, 1959. 392 pp. 64s.

This is an introductory college text-book of astronomy, written by the Field Memorial Professor of Astronomy at Williams College, Mass., U.S.A. The author is evidently a teacher of great experience, bent on prompting in the student "the broadening of intellectual horizons which should result from even a brief study of astronomy". The treatment is almost wholly descriptive, only a few mathematical formulae being introduced. In the latter connection the author states that he has "used English units of measurement freely throughout the book, on the grounds that very few American students can think in terms of the metric system". Such a decision has, however, proved not without its difficulties; for, while velocities are given in miles per second, magnetic fields have to be stated in webers per square meter!

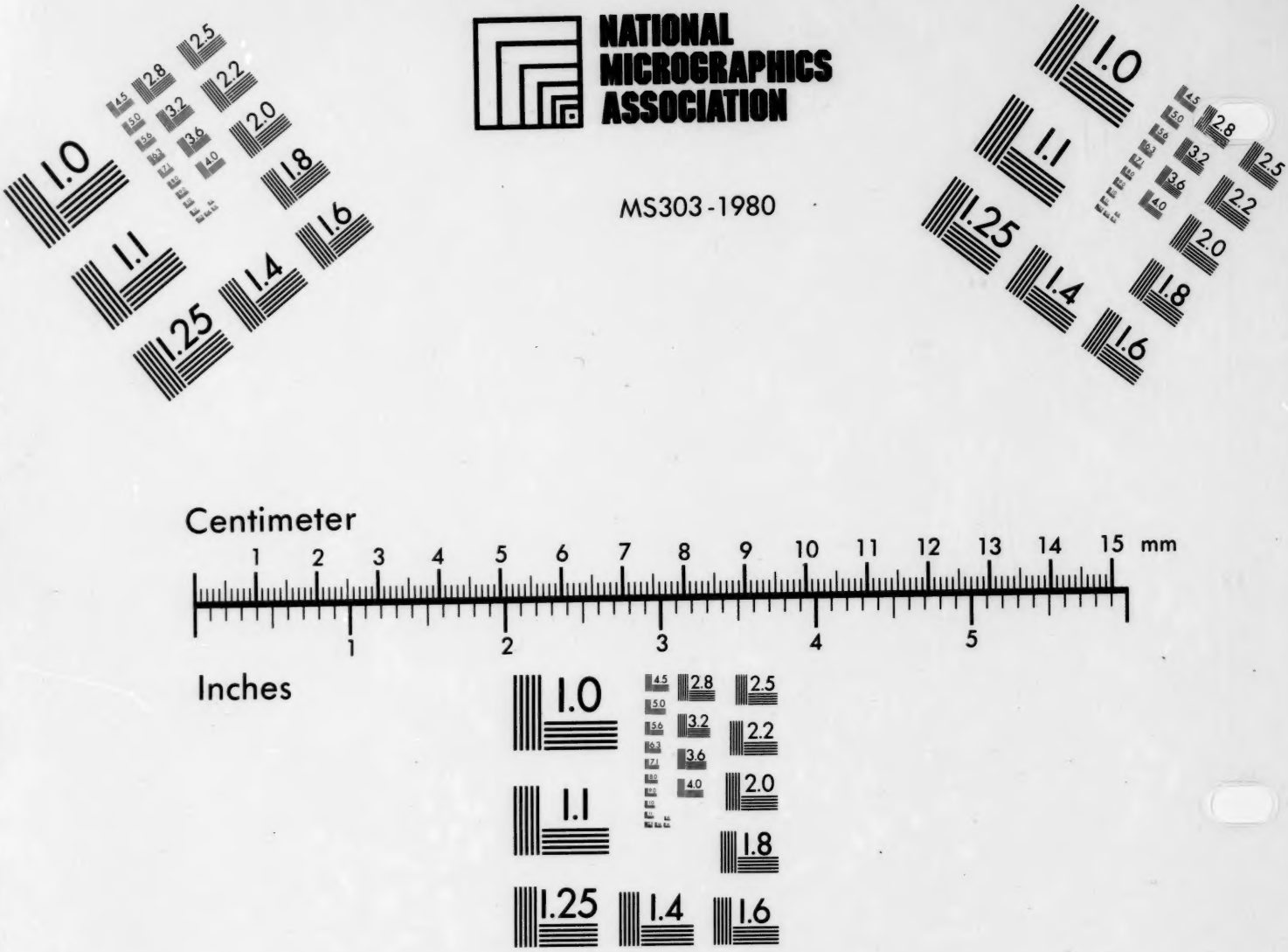
Yet the text is lucidly graphic; and the illustrations—which are to be found on almost every alternate page—are excellent in every way. The fourteen chapters are headed as follows: The Tools of the Trade; The Life Story of a Star; Stellar Distances and Motions; The Sun; Binary Stars; Intrinsic Variables; Galaxies; Our Galaxy; The Characteristics of the Solar System; The Characteristics of the Planets; The Earth; The Earth's Satellites; Comets and Meteors; and The Origin of the Solar System. A welcome acknowledgement of recent events is to be found in the fact that the chapter on The Earth's Satellites includes information about the first seven artificial earth satellites; while, among the test examination questions at the end of this chapter, there is one which runs: "What is the characteristic which determines how long an artificial satellite will remain in orbit? Give the reason for your answer". Thus does modern space research invade the lecture-room.

Unlike some astronomical authors, Professor MEHLIN does not fight a kind of rear-guard action against radio-astronomy. Instead, he welcomes it.

E. V. APPLETON

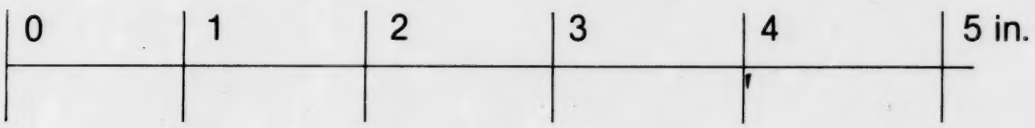
VOL. 16 1959

L.
6
59


NATIONAL
MICROGRAPHICS
ASSOCIATION
MS303-1980
Centimeter
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 mm
1 2 3 4 5
Inches

PERGAMON PRESS INC.

MICROFORMS INTERNATIONAL
MARKETING CORPORATION
FAIRVIEW PARK, ELMSFORD, NEW YORK 10523



PRODUCED AT

**MIMC MICROFORMS LABORATORY
ELMSFORD, NEW YORK**